DOCUMENT RESUME

ED 105 731 FL 006 656

AUTHOR Yaeger, Malcah
TITLE Speaking Syles: Some Phonetic Realizations and their Significance; Pennsylvania Working Papers on Linguistic Change and Variation, Vol. 1, No. 1.
PUB DATE 74
NOTE 63p.

EDRS PRICE MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS *Language Patterns; Language Styles; Language Usage; *Language Variation; Phonetics; Phonology; *Pronunciation; Social Factors; Sociolinguistics; *Speech; Standard Spoken Usage; *Vowels

ABSTRACT

This is an analysis of the stylistic variation in English of a wide range of speakers. A main point is that in analyzing speech styles it is important to elecit more than one style from each speaker, as both interview conditions and other factors may influence style. The paper presents some concrete examples of two major stylistic divisions, spontaneous and controlled, and posits reasons for the stylistic variations. From this evidence, psychoacoustic and psycho-sociolinguistic generalizations are stated. An analysis of acousitic data demonstrates that an understanding of speech production in more than one style facilitates understanding of style variation. Spectrographic measurements mapped the speaker's vowel systems as each speaker was tested as to both spontaneous and controlled speech. Attention especially focused on vowels in a state of flux, as these provide variation. It was concluded that social as well as phonetic factors affect sylistic variation; that variation in stylistic patterns is not limited to only one social or cultural group; and that this type of multistylistic analysis can provide useful information to the dialectologist and the psychoacoustician as well as the phonetician and sociolinguist. (Author/AM)

ED105731

PENNSYLVANIA WORKING PAPERS
ON LINGUISTIC CHANGE AND VARIATION

Vol. I, No. 1

SPEAKING STYLE: SOME PHONETIC REALIZATIONS AND THEIR SIGNIFICANCE

Malcah Yaeger
University of Pennsylvania

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE-
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY.

FL006656

Research supported by
National Science Foundation
Grant GS-36382X

Distributed by
U. S. Regional Survey
204 North 35th Street
Philadelphia PA 19104

PENNSYLVANIA WORKING PAPERS
ON LINGUISTIC CHANGE AND VARIATION

Vol. I, No. 1

SPEAKING STYLE: SOME PHONETIC REALIZATIONS AND THEIR SIGNIFICANCE

Malcah Yaeger*

Linguists have long been aware that the interviewer and the interview situation strongly influence a speaker's style, which in turn strongly influences the data which they gather.[1] Labov's publication of THE SOCIAL STRATIFICATION OF ENGLISH IN NEW YORK CITY (1966) explored the possibilities of collecting and studying style variation systematically, with a strong emphasis on the value of the most casual or vernacular style. But it remains true today that many linguistic studies are based primarily or even totally on one style alone, that being the speaker's most formal style. Socially oriented dialectologists such as Labov, Trudgill, Wolfram, and Shuy, convinced of the necessity of exploring a range of styles in a dialect

*Presently located at the University of California, Irvine.

analysis, have been collecting data with at least two or three degrees of formality. In addition, there is interest in how some parameters of speech reflect style shifting (Shockey 1973).

In this paper I will present some concrete examples of two major stylistic divisions, and posit some reasons why the styles differ for at least the examples cited. From this evidence, specific psychoacoustic and psycho- and socio-linguistic generalizations can be stated: I hope to show the importance of eliciting more than one style from any one speaker. An adequate understanding of the variation inherent in any one style is greatly facilitated by an understanding of speech production in more than one style. In demonstrating this, I will use the analysis of acoustic data which show variation ostensibly too fine for the untrained speaker to hear; yet, since this variation is consistent across speakers, the results of such a multistyle analysis can be useful to the dialectologist and psychoacoustician alike.

Spectrographic analysis has been applied to sociolinguistic data of this type in recent years, allowing us to make much more precise statements about style. Published results appear in Labov 1972a and in Labov, Yaeger and Steiner (1972), which will be the source of some of the data to be discussed here. In some cases, as many as five styles have been utilized, organized along the following

dimension (Labov 1966):

Vernacular---Careful---Reading---Word---Minimal
Speech Speech Style Lists Pairs

There is a major break in this series between the two styles on the left and the three on the right. Both the vernacular and careful styles represent the free flow of uncontrolled speech, in that the observer is not dictating directly the words, phrases or constructions that are used by the subject. The vernacular emerges at points in the interview where the subject is the most deeply involved, emotionally and conceptually, in the events he is describing (excited speech); or at moments where he is not at all involved and where the constraints of the interview situation do not seem to apply (casual speech). Careful speech is the main body of most interviews, in which the subject is answering questions put by the interviewer: it is not defined by any absolute criteria, but relatively to the vernacular forms. Spectrographic analyses of vowels used in these styles show that in most of the communities studied, careful speech differs from casual or vernacular speech by the occurrence of discrete corrections from vernacular style to formal norms, rather than an over-all shift of the phonetic positions of all elements. It has therefore been possible to use both vernacular and careful speech as a single body of data for

our studies of the evolution of vowels, eliminating the most obvious corrections from careful speech. Both of these will be grouped here as "spontaneous style," and opposed to the "controlled style" of reading style, word lists, minimal pairs and other formal tests where the speaker's attention is directed to speech itself. Controlled speech is the normal form used in most phonetic research; spontaneous speech has been the basis for sociolinguistic research which studies the evolution of language in use.

There are two objections which have sometimes been advanced against the proposal that we use not only the controlled style, but the spontaneous style as well. It is obvious that the controlled style can give us the exact range of phonological environments which we suspect to be relevant to our study, with the minimal possible waste of time for both the informant and the researcher. And of course a controlled style can be elicited in a controlled environment so that optimal sound quality can be obtained. The claim is that the extra time needed to elicit a spontaneous interview is very cumbersome, and the burden of proof is left on us to establish that it is in fact absolutely necessary.

Furthermore, phoneticians have long been of the opinion that in any kind of speech in context--even read sentences--the vocalic units never reach their ideal target positions. This paper will attempt to show that when more than one

style of speech is examined for a given speaker, it becomes evident that the speaker aims at more than one target position, i.e., that his target for vowels in spontaneous style is not merely a less clearly articulated, controlled style target, but a target in its own right, for at least those phonological units which are sociolinguistic variables. The difference between these two norms cannot be attributed to a failure to reach the target in contexted speech: in many instances, the controlled style gives us the more central, and therefore the more articulatorily neutral vowel position (Lindblom and Sundberg 1971), while spontaneous speech gives us the more extreme target on the periphery of the vowel space. We therefore feel assured that the difference between the targets is characteristic of the linguistic system, and not due to an impediment in the physical realization of that system. The social significance of these alternate norms has been investigated in studies of the subjective evaluation of linguistic differences through "matched-guise" and "subjective reaction" tests (Lambert 1972; Labov 1966; Labov, Cohen, Robins and Lewis 1968; Giles 1971). Acoustic analyses of the stylistic norms for any given speaker may allow us to study finer variables and locate their place in the linguistic system more precisely.

1. The data and methods of analysis

The data for this paper are drawn from spectrographic analyses of English dialects, carried out in quantitative studies of sound change in progress supported by the National Science Foundation (reported in Labov, Yaeger and Steiner 1972 [hereafter LYS]).[2] I have made a selection of tapes from widely different areas and with a wide range of dialect variables which reveal variation conditioned by style as well as by phonetic environment. By doing so, I hope to demonstrate that the stylistic patterns and their significance for various linguistic interests are not limited to any one social group, but are general to all cultural groups studied to date.

Each speaker studied was chosen for his representative position in a group of speakers of his region, age and class. His vowel system was mapped by spectrographic measurements of 80 to 200 stressed nuclei; details of the methods used are given in LYS, Chapter 2. Special attention was paid to those phonological units which were known to be in a state of change within the dialect.[3]

The phonological units are the vowels which show distinct distributions in word classes in the dialect being studied. The number of vowels measured in each class varied according to whether or not there was evidence of change in progress; primarily, differential distribution

across age levels. The basic data are derived from spontaneous speech. At the same time, for each of the individuals whose speech is discussed here a parallel selection of "controlled" items has also been collected from a word list administered at the end of the interview.

The fully stressed vowels were measured at the points of inflection, minimizing non-dialect-related coarticulatory influence.[4] For the analysis being presented only the first two formants are considered. The tokens of each type are displayed on a doubly linear two-formant grid with F_1 on the vertical axis (with lowest F_1 at the top) and F_2 on the horizontal axis. Such a chart approximates the diagrams of articulatory vowel space.

The present analysis contrasts style charts for the speakers who reveal most clearly the typical pattern of style shifting.

2. Stable vowels

Before we broach the problem revealed by the changes which can occur as the speaker switches styles, it is necessary to contrast such variation with its absence. For any given speaker, the majority of his vowels will usually be stable across styles. These vowels are stable in a more general sense: they are not realized differently by people

of different ages and different social levels within the speech community (Gumperz 1972; Hymes 1968; Labov 1972a: 107, 120ff., 158).

Usually this stability is reflected in the speech of any one subject in the following three ways:

1. A narrow range of variation will be evident among the tokens of the unit in spontaneous style.

2. A very limited range of phonetic conditioning will be evident from the data. (Usually only the most widely recognized of coarticulatory constraints will have any noteworthy effect on the vowel's position in phonetic space.)

3. The target which appears to be aimed at by the speaker does not vary from one style to the next. That is, his word list or minimal pair tokens will fit within the target area revealed by our analysis of the spontaneous style. These vowels are "said the same" whether or not they are a focus of attention. In all diagrams which follow, the unfilled representations on the grid were measured from tokens which appeared in the spontaneous section of the interview, while the hatched representations--whether the ellipses which surround a group of tokens, or the tokens themselves--were plotted from the measurements of the controlled style of the interview. Examples of such stable units can be seen in Figures 1 and 2. Figure 1 shows several vowel units measured in the speech of Eddie G., a

15-year-old Glasgow youth. In his dialect, the "short a" of hat, hand, and the "broad a" of father, hard, calm have not split. There is no phonetic conditioning discernible from the tokens analyzed, and the same target has clearly been maintained for the controlled speech style.

Figure 2 is drawn from the speech of a 16-year-old Philadelphia girl, Jane R. For the moment we will focus on the stable /uw/ before /l/ which can be found in the upper right hand corner of her chart. Jane's tokens of such items as school, pool or other examples of /uwl/ maintain the same formant positions in word lists as in the spontaneous running conversation on which the body of the analysis is based.

In our analyses we have found quite generally that units not in a state of flux within a dialect are pronounced the same whether they are a focus of attention or not: whether the token is measured from spontaneous or from controlled style, it falls within the same area of phonetic space. This is all the more noteworthy since a unit not in the process of dialect flux will have a much narrower range of expectable phonetic values than will a unit presently in a state of change (LYS, Chapter 2). Examples of such stable vowels form a clear contrast with examples of vowels which are not stable, but which are either in the process of change or which still maintain a style contrast in the dialect although the process of change appears to have been completed.

3. Vowels in a state of flux

The body of this paper will be devoted not to the nonchanging, stable vowels, but to vowels with two (or more) target positions which can be shown by our analyses to be in a state of flux. Such vowels can be observed to show the following characteristics:

A. Usually while a vowel is in a state of flux within a dialect, a broader range of articulations will be produced in a spontaneous style than is common for a stable vowel; this is reasonable considering that speakers of other ages or even the speaker himself in a different social setting will use another articulation which varies more or less radically from the one used spontaneously.

B. Consequently, there is more range within which a phonetic conditioning can occur. This will be evident from many of the charts which follow.

C. In addition, the speaker's target for a given token in his controlled style will usually vary from that of his spontaneous style tokens of the same vowel unit. The vowels which are revealed by the results of our analyses of these speakers, their age-mates, and other speakers who come from the same social community, to be in a state of flux, are termed "variables" (LYS, p. 39).

4. Limitations of controlled style

4.1. One limitation of controlled speech as a total data base: The unreliability of introspection. In studying the styles of any one speaker, it appears that the first type of variation involves a general problem of research: an aspect of the uncertainty principle.[5] The very fact that a process is being observed can tend to alter the process. I would like to focus on two potential reasons for this:

A. The self-conscious scrutiny of a variable--or even a non-varying linguistic item--will tend to alter the speaker's perceptions of what he does ordinarily.

B. Another constraint, subsumed under the first but referred to as the 'interviewer effect' by Labov, or 'priming' by Wang, is that the method of observing--in this case the interview form and the interviewer's own relationship (or more commonly lack of relationship) with the speaker--will also tend to alter the results in many ways (systematic or otherwise) from what the speaker will ordinarily do.

Figure 3 is indicative of the type of deformations in the speaker's output which can be revealed from merely listening to her word list. The spontaneous section is close to the vernacular ideal defined by Labov (1972b, Chapter 3), since the subject, Jane R., was being interviewed by a friend and has a strongly expressive personal

style. At the end of the interview, she was asked to read an extensive word list concentrating variables of Philadelphia speech. We can determine from the extralinguistic cues that there is a minimum of correction: she is sincerely trying to use as close to a "spontaneous" style as possible in her rendition of the word list tokens.

The Philadelphia pattern for the (æh) variable is an intricate set of conditions (LYS 3.2.3); it is similar in general outline to the rules operating in New York, Baltimore and other cities in the mid-Atlantic area. On the basis of our own work, as well as impressionistic studies done earlier (Ferguson 1968), the "short a" which is completely stable in Eddie G.'s speech has on the Atlantic coast split into two classes, lax /æ/ and tense (æh), a variable which shifts radically to a higher position. Older and younger speakers show small differences in the assignment of the tense and lax categories, but differ radically in the degree of raising of the tense vowel. The conditioning for this tensing rule is made up of a variety of phonetic and lexical conditions in Philadelphia. The exact raising of the tensed member cannot be determined from a single speaker's spontaneous speech even if it is very extensive. The word list is resorted to because all of the relevant items to determine the stage of (æh) movement cannot be found in spontaneous speech. We might hope to derive two

kinds of information from a word list: which words have been moved from the lax /æ/ class to the tense (æh) class, and where the phonetic sub-classes of the (æh) have moved to in the general raising process.

With all Jane's good intentions of responding spontaneously in her word list, she falters at just those points which are most variable and consequently most critical to our analysis. Jane is in the vanguard of her community's linguistic change: consequently, her spontaneous style is more advanced than that of most people she hears. Her sporadic target shifting is probably more attributable to the variable character of the speech she hears than to socially controlled self-correction. Two potential targets exist for one vowel-unit when that vowel is undergoing or has undergone a process of change. When a speaker cannot even decide between these targets when he is asked to do so, it is likely that the change is still in progress. Consequently, as we shall see, Jane's general linguistic system is revealed in her speech during the interview, while only individually deviant performance is brought out in her responses to the word list presented to her. For example, at two different points in the word list she is asked to give her response to Jan and path. For each of these items she provides two different responses. The readings of these words are deviant in two ways: first, because they oscillate between the tense

and lax categories; secondly, the "tense" pronunciations are not by any means as high or as peripheral as the targets for spontaneous speech indicate by the open ellipses in high front position. It is always possible that Jane's spontaneous style includes both of these options, but given other information from the interview it is more likely that in this case at least, Jane's intuitions failed to give her a sure guide of which norm she ordinarily uses for each item. Another common but trivial deviation may be caused by a mere artifact of word list order; thus sad was on first articulation raised in the word list but is normally lax: in this list there is a row of tensed vowels just before sad.[6]

Such deviations from the spontaneous pattern taken singly may appear as trivial "performance" limitations. However, by being so limiting, they effectively prevent us from using word lists to find out how Jane pronounces the unit at issue in a given phonological environment. Since we interviewed Jane to determine where the extreme advanced edge of change is, her word list responses have masked the effects of change which are of central interest to the dialectologist.

Thus far I have illustrated the general principle that if a variable is not in a state of flux, the speaker will probably not vary his pronunciation of the unit in

different styles of speech; if on the other hand the vowel is in movement, an analysis of a range of styles will reveal that, at the very least, oscillation will take place at predictable points (such as Jane's tokens of sad, planet, or Jan). We can show further that the more a variable is stigmatized, the more likely the speaker is to have a correction at least in one style.

4.2. Another limitation of controlled style: Social-labeling. Figure 4 is an example drawn from the treatment of (æh) in New York City, where the social stigmatization of the variable is much heavier than in Philadelphia. Here the subject's speech can be strongly affected by the self-image that he wishes to project to the interviewer, especially after his attention has been drawn to speech. In fact, once the topic of speech is broached, a subject's productions may only poorly reflect what he would normally do. Leon A. was interviewed by a man of roughly his own age in his own home; however, when asked to read a list of words, his production corresponded to his self-conscious impression of the prestige norm for the same vowels.

This type of variation--correction across styles--is more systematic than the sporadic shifting across styles than we observed from Jane's interview.

Figure 4 shows Leon A. to be a member of the New York City speech community, as can be seen from the clear split

between his /æ/ and (æh) tokens' nuclei gleaned from spontaneous speech (in this case the assignment of the categories follows the New York City form of the rule [LYS 3.2]). But when his attention is directed to speech in word lists, the distinction becomes much less clear.

While the stable /æ/ retains the same target positions for all six nuclei of controlled tokens, (æh) is corrected in the direction of the more conservative /æ/ target and in four instances even overlaps that target; only three tokens have maintained the same target position as they have in his spontaneous utterances.

Returning to Jane's Philadelphia speech in Figure 5, we observe another type of shift in controlled speech, with strong phonetic conditioning. In Jane R.'s spontaneous speech, (æh) preceding a fricative other than f (the variable convention (æhF) is used), such as path or past was the most peripheral subset of (æh). But in reading style she corrects so that none of the (æhF) nuclei overlap between the two styles.

The (æh) before nasal (æhN) is not corrected in controlled style, while her (æhF) is. In general, the subclasses before nasals show the most advanced nuclei and are least often corrected in Philadelphia.[7]

5. Advantages of analyzing several styles

5.1. The advantage of controlled style in conjunct with spontaneous: Articulatory competence revealed. Figure 5 shows one example of the fine-grained phonetic conditioning of (æh) for our Philadelphia speaker; many environments effect a more or less radical change in the location of vowel nuclei in phonetic space. This is one of the effects which can be expected on a vowel in a state of flux, even at as late a stage in that change as we know this one to be.

In any analysis, the first thing we attempt to determine is which environments have what kind of effect on the articulation, and which phonetic conditionings can be determined to hold across speakers of at least one age group. If, for example, Jane were the only Philadelphian to raise (æh) further before nasals, we would say that this condition is not even an interesting influence in the Philadelphia dialect. However, for the Philadelphia speakers we have analyzed, Jane's environmental conditionings in the vernacular are typical.[8]

The effects of Jane's, and other young speakers', consistent phonetic conditioning are not dependent on coarticulation, as that is presently understood. The fact that many of these environmental influences have different effects, even at various stages in the change of one community's speech, shows that such differences have a

strong social component. In some cases these dialectal differences might even be considered as a reaction against the more "natural" effect of coarticulation.

The fact that Jane (like other young speakers) has the consistent stylistic ability to perform the minute differentiation in her (æh) followed by a fricative reveals that the human articulatory mechanism is capable of such a narrowly separated targeting for an environment. It also tells us that Jane R. has the stylistic competence to vary systematically between an advanced target for (æhF) to a less advanced target. This is a social style competence because she consistently differentiates the two targets--the spontaneous from the controlled (æhF) target utterances--by differentiating the spontaneous from the controlled tokens of the same variable. Jane's performance is consistent within her own corpus and consistent with the performance of other speakers of similar social background.

5.2. Another advantage of analyzing two styles: Dual targeting reveals psychoacoustic competence. Figure 6 shows another variable in the speech of Jane R. which registers a shift from spontaneous to controlled style: (aw) in house, count, found, etc. For many Philadelphia speakers, the nucleus has risen to upper mid position, overlapping the nuclei of the front upgliding /iy/ and /ey/ as well as the nuclei of (æh). In Figures 7a and 7b, the relationship of

Jane's (æh) to her (aw) is charted for both spontaneous and controlled styles. The relative positions of the two nuclei are maintained in the shift with extraordinary precision. From this we infer:

(1) Both the (æhF) and (aw) nuclei are in a state of flux in the speech of young Philadelphians relative to stable vowels in their system.

(2) There is a second, controlled target, separate from the spontaneous target, for both of these units.

(3) The controlled target is more conservative (because it is nearer the older [æ] target).

A corollary to (2):

(2') The speakers have the psychoacoustic capacity to control the use of two such narrowly separated targets.

The same message can be drawn from other dialect areas and other patterns of style shifting. Figure 8 shows the vowel system of Carol M. from Chicago. While in spontaneous conversation (Figure 3) Jane's (æhF) and (æhN) differ very little, Carol's spontaneous targets (seen in Figure 8a) are more widely separated. Again this gives evidence of (1) a change in progress. However, Figure 8b (as contrasted with Figure 5) shows that (2) the newer target has become the self-conscious target for which the speaker is aiming in all phonetic environments. As opposed to the Philadelphia pattern, this Chicago pattern reveals the opposite

information about the speaker and/or the change in progress: Carol's internalized self-conscious target is the newer one. More phonetic conditioning is revealed in spontaneous speech; furthermore, in controlled speech, the new more unified target is a much closer targeting of the variable about the mean for all the tokens.

5.3. The target and its social significance. Carol thus adapts herself to a single target in controlled speech. This might be taken to indicate that the change is nearer completion for her dialect community. However, unconditioned targeting is not a proof of change completion as can be demonstrated from a second look at Jane's Figure 6. No strong phonetic conditioning is shown by Jane's (aw) in her controlled (or her spontaneous) speech. Nevertheless, in contrast to Figure 6, Figure 9 reveals that for some young Philadelphians the movement of (aw) is hardly begun. Consequently it is not likely that the raising of the (aw) nucleus is a change nearing its completion for the speech community as a whole. Some dialect phonologists might posit here a merger of Philadelphia (æh) and (aw) nuclei; but--granting any psychological validity/reality at all to a phonology--that position must be ultimately untenable.[9] The nuclei of the variables--irrespective of glide--cannot be equated in the informants' minds if the social attitude connected to the changes (as revealed by the social patterns of

production for the respective variables) is not consistent across speakers. It is part of the psycholinguistic competence of the speaker to know whether or not a given unit is stigmatized within his community. Our Philadelphia data indicate that speakers treat (æh) and (aw) very differently in this respect: (aw) shifts as a whole (Figure 6), or not at all (Figure 9), while some conditioned allophones of (æh) are corrected more than others (Figure 5). Given this differential treatment of the two variables, the speaker cannot be considered as having one merged (aw)-(æh) nucleus, as convenient and as elegant as such a solution might appear on the basis of one subcommunity's spontaneous utterances.

In addition, as Labov, Hymes and many others have consistently maintained in recent years, the communities' attitudes toward a change in progress, and their consequent manipulation of their production in a controlled style, are themselves of intrinsic interest. The contrast in attitudes between Leon A. and the two girls being discussed reveals contrasting types of <u>speakers</u> who can come to our attention. The contrast in attitudes between Jane and Carol reveals contrasting <u>community</u> outlooks on the tensing and raising of (æh). Both individual and community patterns are of interest to the sociolinguist.

5.4. Competence in production/perception as revealed by style shift. The psychoacoustic issues are more complex than our treatment has presumed so far. It has been shown (Fant 1967) that vowel perception is potentially less discrete than stop consonant perception and that vowels are more easily perceived in the gradient non-speech mode. Only in clearly distinguished speech perception tests can vowels be analyzed discretely by the speaker.

A further complication is that production of vowels also tends to be more variable even within the utterances of a single style of a single speaker. Nevertheless, V-perception in a speech context for any given vowel will tend to be limited by the community's phonological categories. In addition, at least two factors must be considered. Both of these are evidenced in earlier work (Labov 1966; Trudgill 1972).

(a) Within a phonological vowel unit, if a discriminatory social load is carried by one subvariant over another, the ability to distinguish among these variables will be directly affected.

(b) Across phonological units, social pressures may relieve the discriminatory load, thus reducing the subjects' ability to distinguish between the units.

In either of these cases the results of perception tests will not conform with the expectations of phoneticians

or phonologists: phoneticians might look for a more gradient breakdown, while phonologists expect the discrete units perceived to correspond more exactly with the phonological units.

The social load on a vowel distinction can be minimized favoring non-discrimination across units. This will be revealed as a perceptual merger. Examples of such mergers have been presented in LYS (229ff., section 6) and in Nunberg (1973). These include:

(6.1) /ohr ~ oh/ in the NYC vernacular

(6.2) /a ~ oh/ in central Pennsylvania

(6.4) /uw ~ ow/ in Norwich, England

(6.5) /ay ~ oy/ in Essex, England, in conjunction with Nunberg (1973)

Nunberg (1973) reveals that not only can social perceptions differ from the reality of phonology, but a distinction can be maintained despite perceptual judgments made across several generations: the "merger" of ay/oy was attested as early as the 16th century, but the unacknowledged distinction is still being maintained in spontaneous utterances in Essex today.

The social load on a distinction can be maximized by the testing method. It is possible to elicit discrimination on pair tests in some cases. But in both pseudo-merger and social/phonological correction, the most reliable means

of analyzing speakers' discriminatory capacity is obtained in actual spontaneous speech data. Just as the Essex data reveal that a minor F_2 difference has probably maintained the phonological (ay)/(oy) distinction for 300 years, we have seen that a similar minor F_2 shift can differentiate two styles of speech for conditioned variables of a unit (as in Jane's (æhF) or for an entire unit such as (aw)). Here the unacknowledged load is not phonological but stylistic-sociological. Since this stylistic variation is an accompaniment to a sound change in progress, it is unlikely that an arrested change could be maintained as a style distinctive vowel color for any length of time. Nevertheless, while such a change is in progress, looking at multi-style analyses of the variable unit reveals more about the speaker's ability to discriminate than can be revealed by any perception test yet devised.

Another way of looking at these phenomena is that of Lisker (LYS footnote 21 to section 6), who suggests a split between the "hearing" and "labeling" responses of the subjects. Until such time as a true "hearing" test can be devised, we will be limited in our analyses of perception to what subjects label as being heard, and what subjects produce in a systematically discrete manner. Since much finer systematic distinctions are revealed by analyses of production, it seems advisable to judge speakers'

perception by these rather than by their (apparently) cruder ability to label.

Let us now consider one aspect of the variable processes in somewhat greater detail: correction. It is now obvious that when the subject is made self-conscious about a change progressing in his speech he will "correct." As early as the Martha's Vineyard study (1963), Labov documented that a speaker will correct in the direction of the image of himself which he wishes to project to other people. If this is in fact the case, the interactional component is bound to become stronger as the speaker's self-consciousness about his speech increases.

Jane and Leon both correct in the direction of the older standard as they read their word lists. On the other hand, Bill P. of Duncannon, Pa. (LYS:6.2) corrects toward the newer norm in his word list production [see Fig. 10, from LYS Fig. 6-4].

While this older rural speaker maintains his own norm of /a/ ≠ /oh/ in his own non-self-conscious interview, in the controlled speech of the word list, he chooses to be considered as a speaker of the newer regional dialect, with [ɒ] for merger /a,oh/, which however did not even begin to develop until <u>his</u> own dialect had been formed.

The problem of understanding the ways in which the speaker will attempt to conform to a norm--whether it be an older form of his own dialect, or his understanding of

the interviewer's prestige norm, or some third option--can become increasingly complex. Unraveling the motivation behind the aberrations which occur in formal speech provides an interesting puzzle; but in the course of a dialect analysis in which we have not determined the spontaneous norm of the speaker, the use of word lists in a social dialect study can be more confusing than revealing, since its relationship to spontaneous style is far from obvious in any given situation.

In the following analyses from a survey of farmers from Southwest Utah, the problem being explored was the apparent flip-flop of (ahr)--as in far and part--and (ɔhr)--as in for and port. It is traditionally reported that older rural speakers reverse these two word classes. An impressionistic study by Cook (1969) reported considerable overlap, and threw doubt on the existence of a true reversal (see summary in Labov 1972b:281). But our spectrographic studies of rural speakers from the same area show that the "ahr" words are backed and raised, and the (ɔhr) words lowered. For older speakers there is a reversal with some overlap, and for younger speakers there can be a complete reversal. Figure 11 shows this pattern in the spontaneous speech of a Kanab rancher (11a) and his daughter (11b) who were interviewed together.

Figure 12 shows the analysis of spontaneous productions of (ahr) and (ɔhr) in interviews with three older men from

the area. Although in each case there is a great deal of overlap, in all cases (ahr) has already risen to a higher mean position than (ɔhr). This distribution is corroborated by the reversal of the means, marked with an X for each variable.

Figures 13a and 13c show that the two younger men (80 and 69) in fact do reverse and correct this stigmatized flipflop in controlled style. While there is still a great deal of overlap between the tokens of (ahr) and (ɔhr), (ahr) is now found in the lower position. The eldest speaker, 85, does not correct in this way, but retains his local linguistic identity.

Given the degree of overlap in Figures 12 and 13, we would have to characterize the three older speakers as showing a merger of (ahr) and (ɔhr) in spontaneous speech, while younger speakers like the 50-year-old daughter in Figure 11b have carried the reversal to completion.

A problem in delineation of the variables arises. Data is limited: though the interviews are from one to two hours long, there are still insufficient tokens of the relevant variables for them to tell us much about the environmental phonetic/phonological conditioning. Nor, as we have seen, is there hope of determining this conditioning from the word lists elicited from the speakers. The reversal of (ahr) and (ɔhr) has become a social stereotype (Labov 1972b:314) for

rural speech, as the subjects freely admit: to whatever degree an interviewee monitors his speech under the effect of observation from an outside source, he will correct toward what he considers the standard to be or toward the local dialect. Thus a true picture of the developing linguistic change across age groups is impossible on the basis of controlled speech. Even more difficult, of course, is any hypothesizing about the relative influences of various phonological environments on the variables.

Among the older speakers whose apparent merger of the variables appears on Figure 13, only Claude C., the oldest speaker, maintains this merger; in fact he appears to correct toward the stereotype while the other two men interviewed have obviously attempted to correct toward the standard, reversing their spontaneous pattern. While there is still a great deal of overlap, (ahr) for these speakers is found in the lower position closer to the standard in controlled style.

The tendency to correct toward a newer norm is more prevalent in a society with high group pride (Garvin and Mathiot 1960); this appears to be the case in England. In fact a recent article by Peter Trudgill (1972) concerning social prestige in his community also argues that the--as he terms it--covert prestige of the working class norm is much stronger in England. Trudgill very nicely documents this

point for an entire community, with the aid of self-report tests performed in Norwich. He is also of the opinion (cf. especially pp. 184, 188) that this opposite correction is generally indicative of the British working class consciousness as opposed to the U. S. "conspicuous lack of . . . class consciousness. . . which is one of the most important contrasts between the American and European systems of stratification." Since his evidence corroborates other reports and our acoustic evidence quantitatively corroborates his impressionistic data, it follows that the community's pattern for deciding what is prestigeful must be considered when forming hypotheses as to just what deformation of the actual usage is occurring in any given style of speech. The different phonological profile revealed by the section of the interview conducted in controlled style is indicative of the pattern of targets which the speaker wishes to emulate. Only the accompanying sociological data and insight can reveal what norm it is that the speaker is emulating, whether the overt middle class norm, or, as in the case cited, a differing, more local working class norm.

In an exact parallel with LYS data previously cited from Western Pennsylvania, we can turn to data from the community analyzed in detail by Trudgill. In Norwich such words as beer and bear are merged in working class

speech. See Figure 14a for Tony T.'s (ihr)~(ehr), which corresponds to Trudgill's (er) variable. In the speech of David B. (Figure 14b), these word classes are clearly distinct. Yet David's (ihr)~(ehr) as elicited in the word list (Figure 15) document again that even lower middle class speakers are aiming at a merged Norwich target. Not only is David clearly merging the two units in controlled style; in his self-report of what he and other Norwich residents say, David "under-reports" his own production. (That is, in Trudgill's term, he under-reports his output as defined by the middle class prestige norm.) We could infer that he would do so by perusing the measurements from his own interview, even if we did not have Trudgill's corroborating evidence, which tells us that many male speakers of Norwich will similarly "under-report" themselves for this very reason.

Thus, for some dialect speakers (like Bill P. or Carol M.)--and for some communities' dialect attitudes (like those of Norwich)--it is the word list which can reveal most clearly what the emerging dialect pattern is most likely to be. An even more dramatic example of the interviewee correcting toward the newer target in controlled style can be brought from Scottish data. We will be concerned here with the realization of the class of Middle English ū words, in <u>house</u>, <u>out</u>, <u>found</u>, etc. In nonstandard

Scots dialects, the diphthongization and lowering of the nucleus characteristic of the Great Vowel Shift did not take place; but in the modified Scots-English which serves as a Glasgow standard, this diphthongization is found and the nucleus is lowered to mid or upper mid position, rather fronted: [ə˓u]. In the meantime the vernacular form has been fronted along with the reflexes of M. E. ō and ŭ. As Figure 16 shows, the nuclei of all three vowels are found in roughly the same position for speakers of various ages: (Figure 16a) Mary W., 79; (16b) Jim McN., 21, and (16c) Totty R., 14. Though Totty is the youngest of the three, he does not maintain this pattern consistently in spontaneous style.

Totty was interviewed along with his best friend Eddie. The fact that two peers were interviewed together is significant, especially for younger people, because in these peer interviews the subjects become involved in what they are discussing, and will talk both more naturally and more volubly. In addition there is the basic peer pressure not to affect a "posh" style.[10] The chance of hearing a strongly affected dialectal pronunciation is consequently minimized. Yet while Totty's ō and ŭ are consistently fronted, with the direction of the glide in specifiable instances determined by the following consonant, the word class of house, about, now, etc. is variable. We will refer to this

variable as (ū), after the historical form which is the most conservative of the three norms involved.

Figure 17 shows the over-all distribution of Totty's (ū) forms. From his spontaneous speech, we isolated and measured twelve fully stressed tokens of (ū): four followed the new ($ū_3$) norm with a nucleus at [ü], two are not quite as fronted, two are found in the conservative high back ($ū_1$) position at [u], one of which glides <u>toward</u> the newer target, and two more glide from the new to the old target. Two are found in the area characteristic of the standard diphthong ($ū_2$). In contrast, the five tokens of (ū) from the word list <u>all</u> follow the new norm.

This can be seen more clearly in Figure 18, which shows all Totty's spontaneous (ū) nuclei with the direction of the glide shown and contrasts them with the group of controlled (ū) forms. Not only do the word list tokens cluster most closely, but they effectively demonstrate that when asked to reflect on his speech, Totty projects an image as a speaker of the local Glasgow dialect, bringing the nucleus of (ū) to the [ü] position.

The fact that David and Totty correct toward the working class norm in their controlled speech tokens tells the sociolinguist something about these boys as individuals. In a larger context it tells us something about the speech communities which the boys are members of. The

fact that for both Totty and David there is one style which reflects a single consistent target for the variable, and another style with a systematically differing target, allows us, here too, to be confident in our judgment that the boys have a finer sense of phonetics than could be clearly elicited from their own self-conscious perceptions.

Such patterns as those revealed in Totty's and David's interviews are likely to be found in any urban dialect at any stage and, in fact, reveal a case of what Labov has termed "inherent variability." An analysis of any given speaker's total range of variability would consequently be quite revealing. But for our present focus, I merely wished to make clear that given, e.g., Totty's much greater degree of variability for (ʊ), and given the fact that while the range of variability is broad, the fact of differential variability is present for all Scots speakers analyzed to date, as well as the cases actually described in the literature (Grant and Dixon 1921; Ellis 1889; Scottish National Dictionary), it is not possible to claim, as one surely would on the basis of word lists alone, that nuclei of Middle English (ū, ō, ŭ) are merged in any psychologically meaningful way. If we posited such a merger, we would ignore many facts which clearly reveal much about the speaker's understanding of his own system.

6. The effects of observation

The Glasgow and Norwich speech communities are two among many which provide data that underline the impossibility of avoiding the interviewer effect, or what Labov has called "the observer's paradox" (1972b:209). Any attempt to ignore this effect will lead to serious distortion. Instead, we can use the variety of stylistic contexts to measure the size and direction of the shifts which are due to the subject's perception of the situation and the norms that govern his behavior.

The linguist interviewing in his own speech community is handicapped because his own dialect marks him as a member of a particular social group and maximizes the shift of working-class speakers towards the local prestige norm which he exemplifies. While the outside interviewer may gain from the socially unmarked character of his dialect (as with our own American accent in Norwich or Glasgow), he lacks the knowledge of local norms and word classes which will allow him to readily understand and interpret the data he obtains. Spontaneous speech styles are less subject to the strong social correction that is applied to a well-established stereotype; in the absence of such correction, controlled speech may sometimes give us a clearer representation of the vernacular norms, but rarely

yields as clear a view of the social and phonological knowledge possessed by the speaker.

We have already observed that merely being forced to a self-conscious attitude toward speech alters the output of the speaker. The ways in which this awareness alters the speech and the consistency with which it alters it are, however, dependent on the cultural community's expectations. This understanding is explicit in the work of the researchers cited. Implicit in their work is the understanding that the relationship of the speaker and his interviewer can alter the base of expectations which the interviewee relies on for deciding his style. It stands to reason that as the interview situation becomes increasingly more formal, the speaker's style will reflect an increasing deviation from his own view of what he thinks he says, and will approximate--to whatever degree the speaker is able--his mediated stereotype of the impression he wishes to leave with the interviewer; whether that is nearer to the standard or to a more local prestige norm will tell us a great deal about the speaker and his community. The controlled speech elicited in the interview will differ more or less radically from spontaneous speech depending on the sociolinguistic community's consciousness of the change in progress, and their attitude toward it; it will also depend on the subject's own consciousness of

and attitude toward himself and the relationship between himself and the interviewer. The changes which a given speaker can perform on his speech in the process of style-shifting reveal the difference in and the direction of linguistic change--but the results also yield a rich body of information on the nature of phonetics and social competence which a speaker possesses.

Footnotes

[1]This paper and the work reported here, prepared in the spring of 1973, form part of a research project on "A Quantitative Study of Linguistic Variation and Change," under National Science Foundation Contract 36382X. I am indebted to the NSF for support on this work, and to the project director William Labov, whose influence will be most evident, as well as to my coworkers on the project--especially David Depue, Gregory Guy, Carolyn Fuller and Laurel Dent for their assistance. I am also indebted to Professors Leigh Lisker, Vicki Fromkin and David Sankoff for their willingness to discuss the paper, and to Peter Trudgill for rewarding discussions of the Norwich data.

[2]The data on which this report is based were gathered by W. Labov and the research group headed by him, first at Columbia University and then at the University of Pennsylvania; tapes are gathered in the field--that being the home of the interviewee or a nearby place frequented by him--his place of work, a park, or a friend's home. The data used will not of course come from all of these sources, but from a selection of the tapes being collected systematically by our own project. These tapes are recorded on Nagra tape recorders at 3 3/4 inches per second. Each interview ranges from one to two 5-inch reels, or from between 45 minutes to 1 1/2 hours. Of course all of the tape does not consist of continuous speech by the interviewee, since that would not simulate the style of natural interaction. In the major part of the interview, the interviewer attempts to overcome the constraints of the question-and-answer situation by involving the subject in conversational interaction. The more formal sections include minimal pair tests, commutation tests, and other types of controlled speech (see below).

These methods have been used and developed by individuals and groups as far flung as P. Trudgill in Norwich (1972), H. Cedergren in Panama City (1973), and the Detroit survey headed by R. Shuy (Shuy, Wolfram and Riley 1967).

[3]This does not in fact assure that all relevant changes-in-progress will always be analyzed sufficiently carefully on first analysis. For example, our first analyses of Scots dialect did not take great note of the range of (ay) as in night or fight. Only after studies of younger speakers had

made it increasingly evident that (ay) for these speakers is as radically advancing a unit as (æh) is for most of the United States dialect groups studied (cf. LYS Chapter 3) did we return to the data and collect a more thorough sampling of the variable in the relevant environments. Most references here to "change in progress" will refer to changes of a given phonological unit in age distribution or changes in "apparent time," supported by auxiliary evidence from earlier observations in real time (Labov 1972b: 163).

[4]For a more precise description of the methods used in measuring vowel formants, see LYS:29.

[5]See Platt (1964) for an enlightening discussion of this issue.

[6]We have all come across this sort on indecision on the part of our subjects. When we do not, that is often the time to be most wary, if we are aware that some dialect feature may be at issue. Labov (1972c:70-71) gives a number of inconsistencies between intuitive reports and forms actually used in speech. On another portion of that questionnaire, directed toward New Englanders, a linguist friend was just as adamant that he had never come across the form at issue: So don't I. Later his younger brother, a silversmith, arrived, and not only demonstrated a native command of the form, but in addition derided his brother for having missed the fact that Aunts A and B, not to speak of his wife's mother, used the forms all the time.

[7]Ohala (1970) and others have documented the articulatory constraints imposed by the lowering of the velic obstruction to the nasal passage. Our studies demonstrate that while ultimately the effect on the preceding vowel is that it rises, there are social dialects for which we have found that not only is the "natural" articulatory influence not present, but quite the contrary. See also Ferguson (1973) and its bibliography.

[8]There are also, as mentioned, purely articulatory limitations on what Jane can do; in our explanations we must be aware of which of the environments can be termed non-social sources of variation in providing explanations. On the other hand, some of the presumed "articulatory" constraints can be determined, on a close analysis of

phonetic conditioning within a given dialect, to be socially variable, proving their potential social variability for any dialect.

[9]The view of psychological reality as discussed here is most clearly allied to that of Chen (1970).

[10]In Totty and Eddie's group, anyone who talks "posh" may be "slogged" for his affectation (Totty R., IV# A456).

References

Cedergren, H. The interplay of social and linguistic factors in Panama. Unpublished Cornell University dissertation, 1973.

Chen, M. "Vowel length variation as a function of the voicing of the consonant environment." PHONETICA 22: 129-159, 1970.

Cook, S. Language change and the emergence of an urban dialect in Utah. Unpublished University of Utah dissertation, 1969.

Ellis, A. J. EARLY ENGLISH PRONUNCIATION. Part 5, 1889. New York: Greenwood Press reprint, 1968.

Fant, G. "Theory of speech analysis: Sound, features, and perception." RIT STL/QPSR 2/3:1-14, 1967.

Ferguson, C. "'Short a' in Philadelphia English." Mimeographed, 1968.

_____. "Nasals as a universal." Paper presented at the Forum Lecture at the Linguistics Institute, Ann Arbor, Michigan, August 1973.

Garvin, P. and M. Mathiot. "The urbanization of the Guarani language--A problem in language and culture." In J. Fishman (Ed.), READINGS IN THE SOCIOLOGY OF LANGUAGE. The Hague: Mouton, 1968.

Giles, H. "Patterns of evaluation to R. P., South Welsh and Somerset accented speech." BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY 10:280-281, 1971.

Grant, W. and J. Dixon. MANUAL OF MODERN SCOTS. Cambridge: University of Cambridge Press, 1921.

Gumperz, J. "The relation of linguistic to social categories." In D. Slobin (Ed.), A FIELD MANUAL FOR CROSS-CULTURAL STUDY OF THE ACQUISITION OF COMMUNICATIVE COMPETENCE. Berkeley: University of California Bookstore, 1972.

Hymes, D. "The ethnography of speaking." In J. Fishman (Ed.), READINGS IN THE SOCIOLOGY OF LANGUAGE. The Hague: Mouton, 1968.

Labov, W. THE SOCIAL STRATIFICATION OF ENGLISH IN NEW YORK CITY. Washington: Center for Applied Linguistics, 1966.

_____. "The internal evolution of linguistic rules." In R. Stockwell and R. Macaulay (Eds.), LINGUISTIC CHANGE AND GENERATIVE THEORY. Bloomington: University of Indiana Press, 1972a. Pp. 101-171.

_____. SOCIOLINGUISTIC PATTERNS. Philadelphia: University of Pennsylvania Press, 1972b.

_____. "Where do grammars stop?" In R. Shuy (Ed.), GEORGETOWN MONOGRAPH ON LANGUAGES AND LINGUISTICS 25, 1972c. Pp. 43-88.

_____, P. Cohen, C. Robins and J. Lewis. A STUDY OF THE NONSTANDARD ENGLISH OF NEGRO AND PUERTO RICAN SPEAKERS IN NEW YORK CITY. Philadelphia: U. S. Regional Survey, 1968.

_____, M. Yaeger and R. Steiner. A QUANTITATIVE STUDY OF SOUND CHANGE IN PROGRESS. Report on NSF Contract GS-3287. Philadelphia: U. S. Regional Survey, 1972.

Lambert, W. E. at al. "Evaluational reactions to spoken languages." In W. E. Lambert, LANGUAGE, PSYCHOLOGY, AND CULTURE. Stanford: Stanford University Press, 1972. Pp. 80-96.

Lindblom, B. and J. Sundberg. "Acoustical consequences of lip, tongue, jaw and larynx movement." JOURNAL OF THE ACOUSTIC SOCIETY OF AMERICA 50:1166-1179, 1971.

Nunberg, G. "The reversal of a merger in 18th century English." Paper presented at Linguistic Society of America Summer Meetings, Ann Arbor, 1973.

Ohala, J. "Aspects of the control and production of speech." UCLA WORKING PAPERS IN PHONETICS 15, 1970.

Platt, J. "Strong inference." SCIENCE 146:347-352, 1964.

SCOTTISH NATIONAL DICTIONARY. Vol. I, Introduction. Edinburgh: Scottish National Dictionary Association Ltd., 1931-1971.

Shockey, L. "Some frequently recurring phonological processes in English." Paper presented at Acoustic Society of America meetings, Boston, April 1973.

Shuy, R., W. Wolfram and W. Riley. A STUDY OF SOCIAL DIALECTS IN DETROIT. Final Report, Project 6-1347. Washington: Office of Education, 1967.

Trudgill, P. "Sex, covert prestige and linguistic change." LANGUAGE IN SOCIETY 1:179-195, 1972.

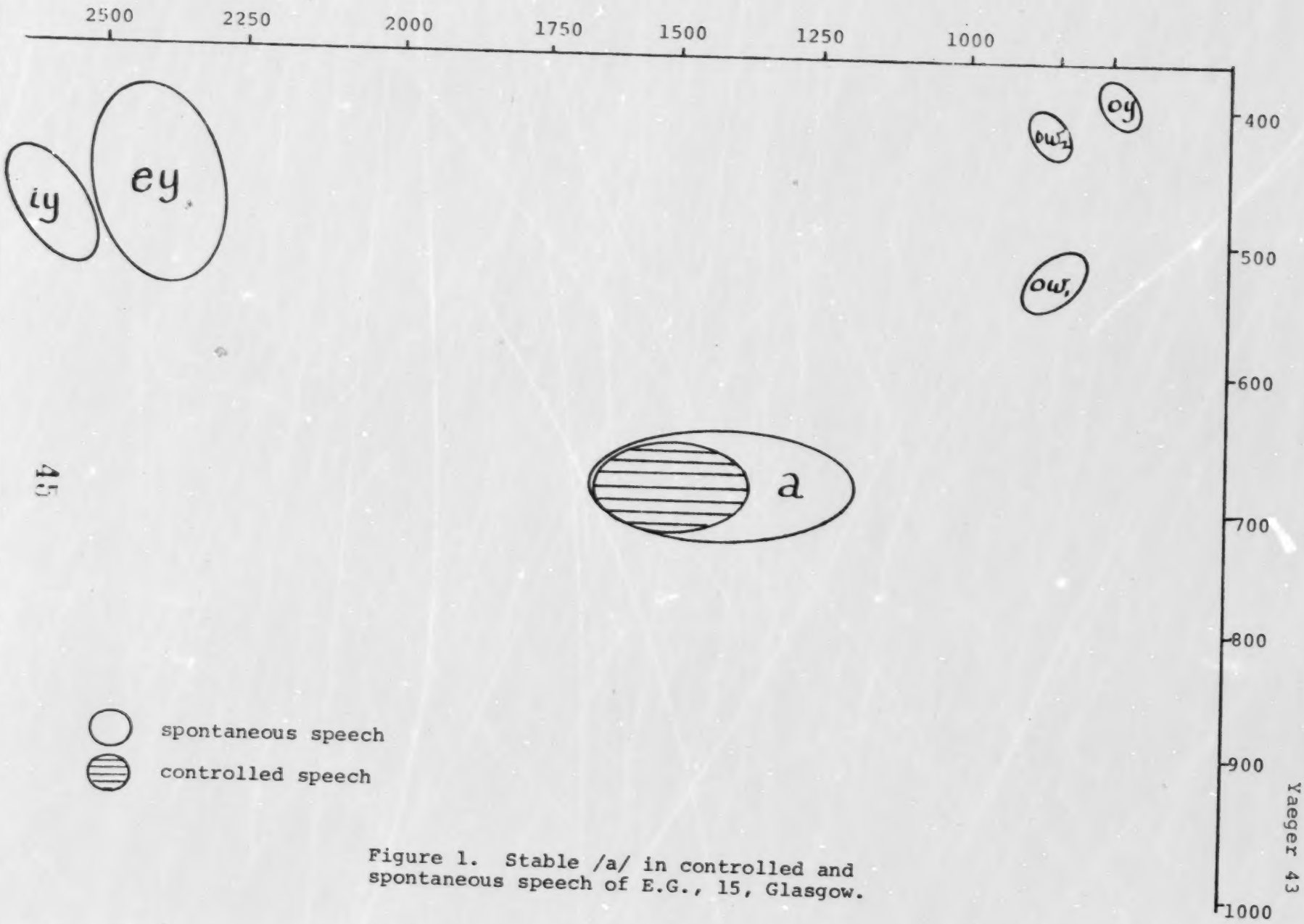


Figure 1. Stable /a/ in controlled and spontaneous speech of E.G., 15, Glasgow.

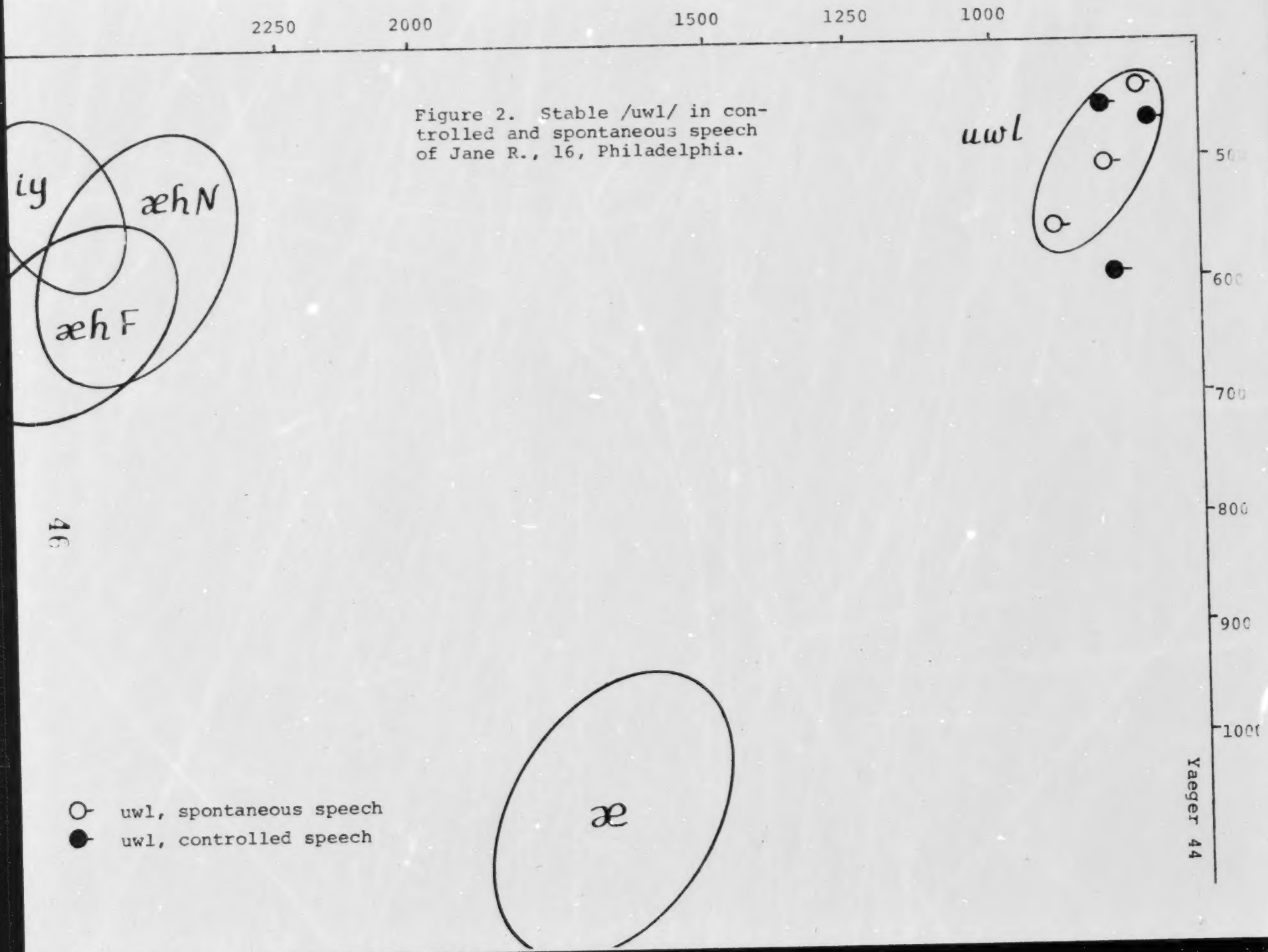


Figure 2. Stable /uwl/ in controlled and spontaneous speech of Jane R., 16, Philadelphia.

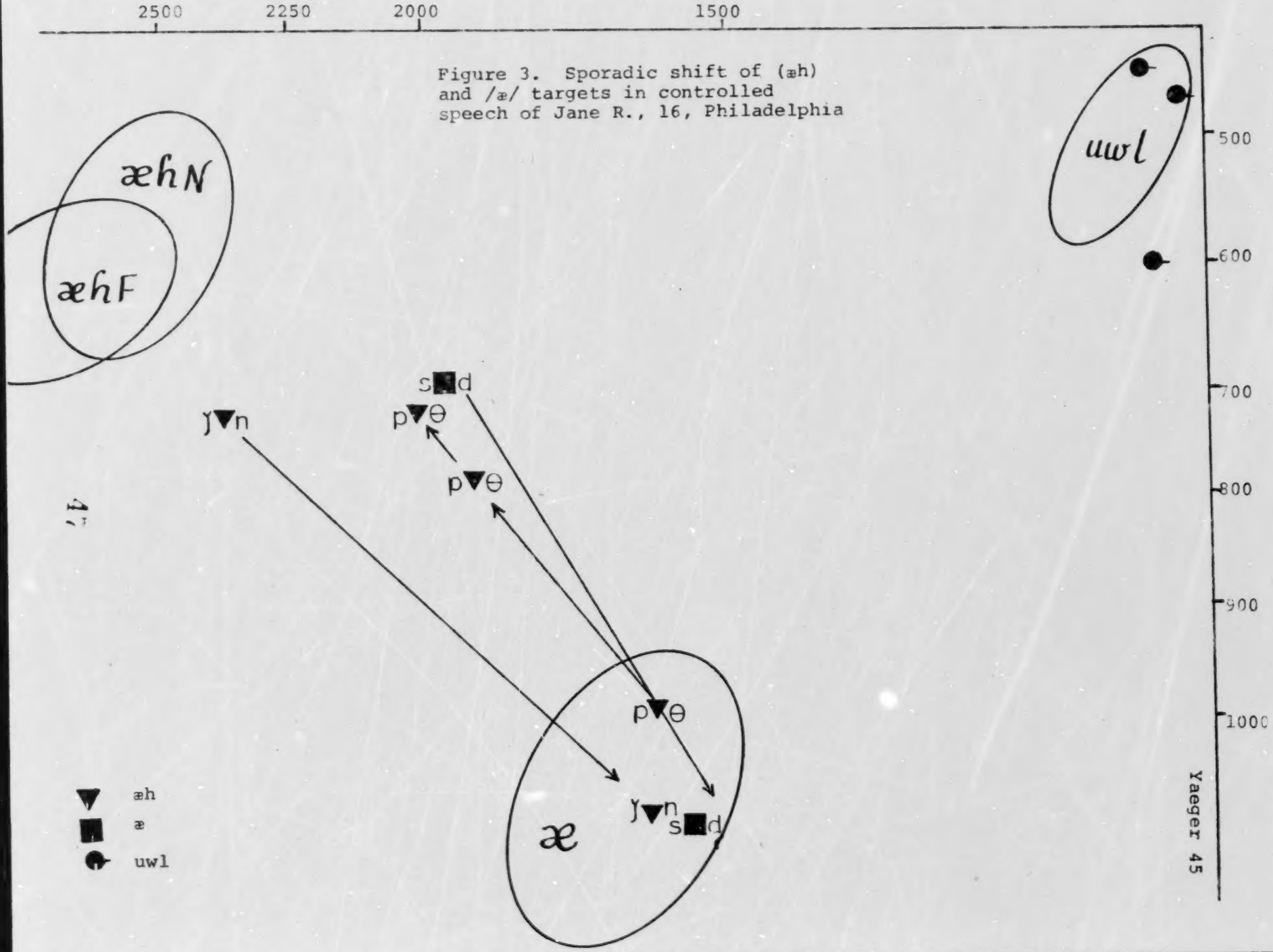
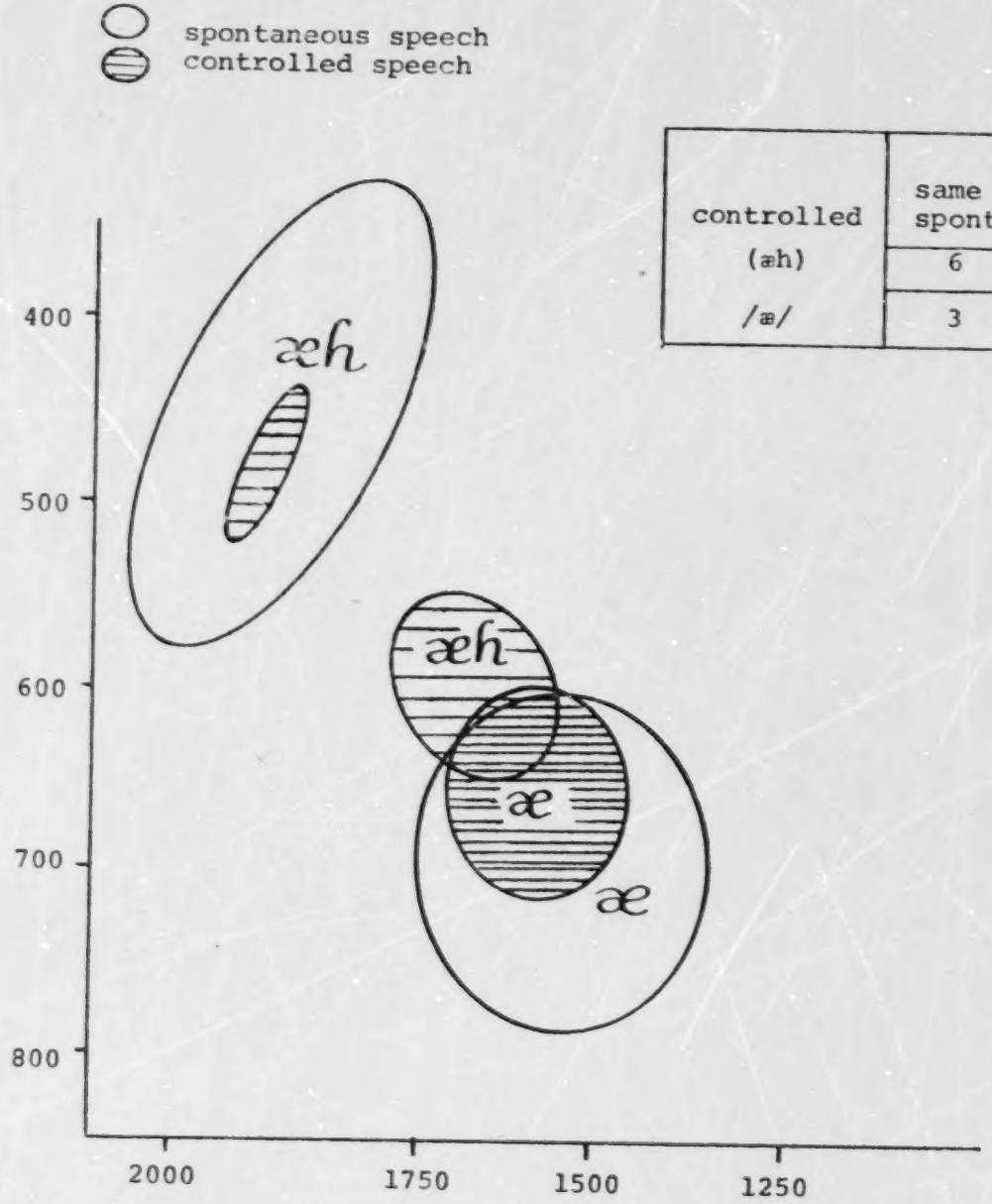


Figure 3. Sporadic shift of (æh) and /æ/ targets in controlled speech of Jane R., 16, Philadelphia



controlled	same as spont.	different from spont.
(æh)	6	0
/æ/	3	11

Figure 4. Irregular correction of (æh) in controlled speech by Leon A., 35, New York City

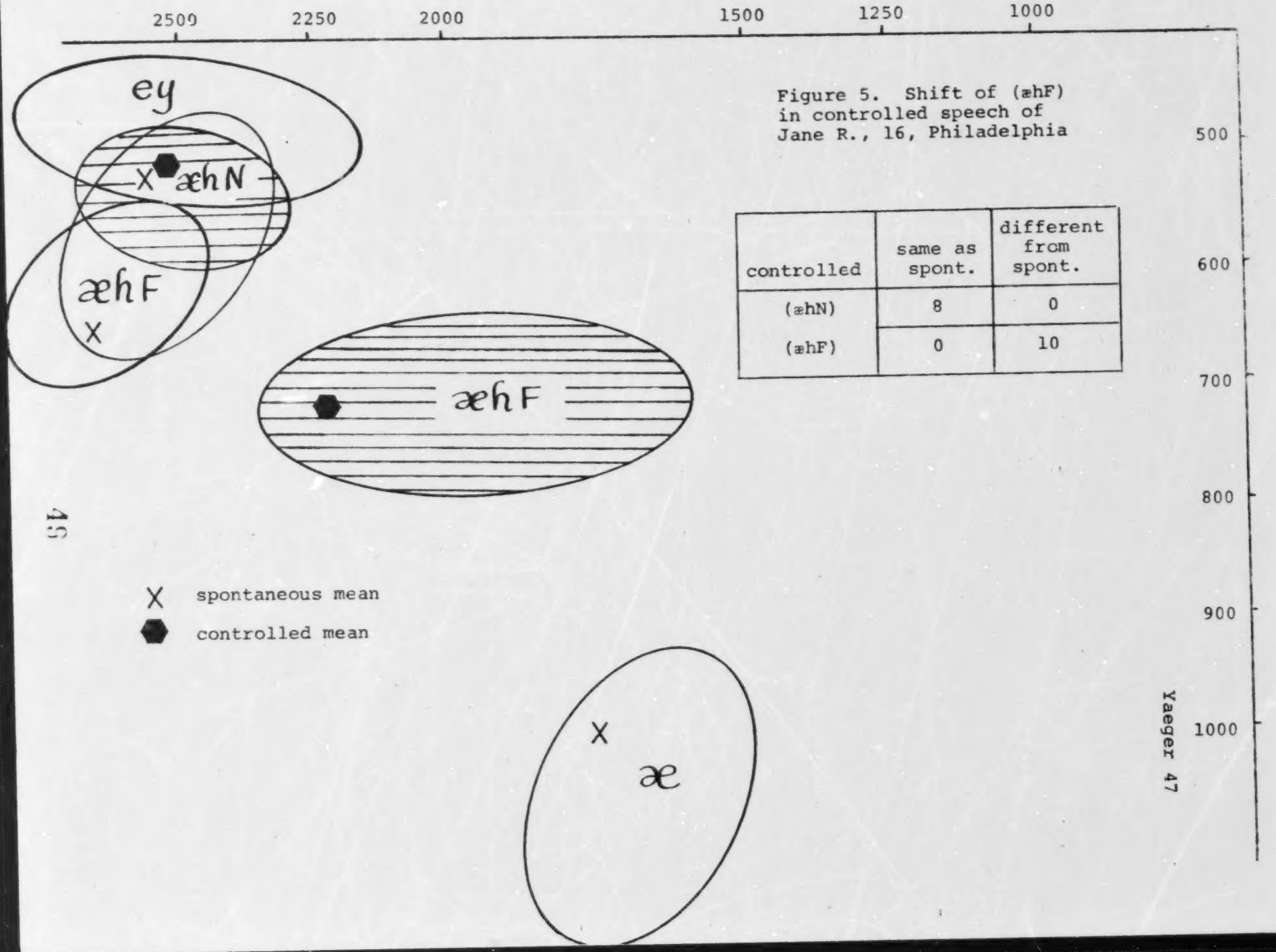
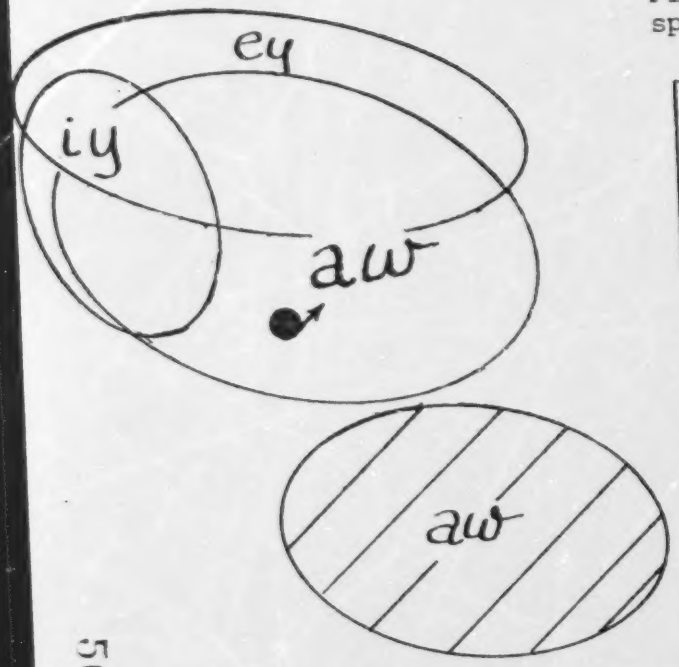


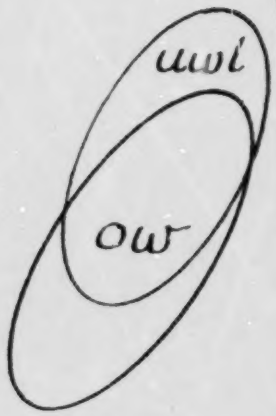
Figure 5. Shift of (æhF) in controlled speech of Jane R., 16, Philadelphia

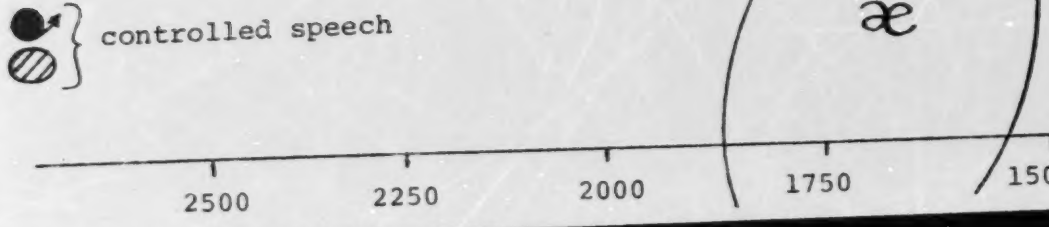
controlled	same as spont.	different from spont.
(æhN)	8	0
(æhF)	0	10

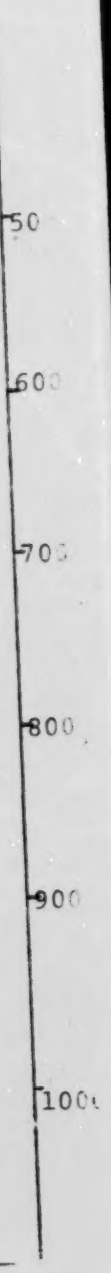
Figure 6. Shift of (aw) in controlled speech of Jane R., 16, Philadelphia.

	same as spont.	different from spont.
controlled (aw)	1	6









controlled speech

Yaeger 48

50

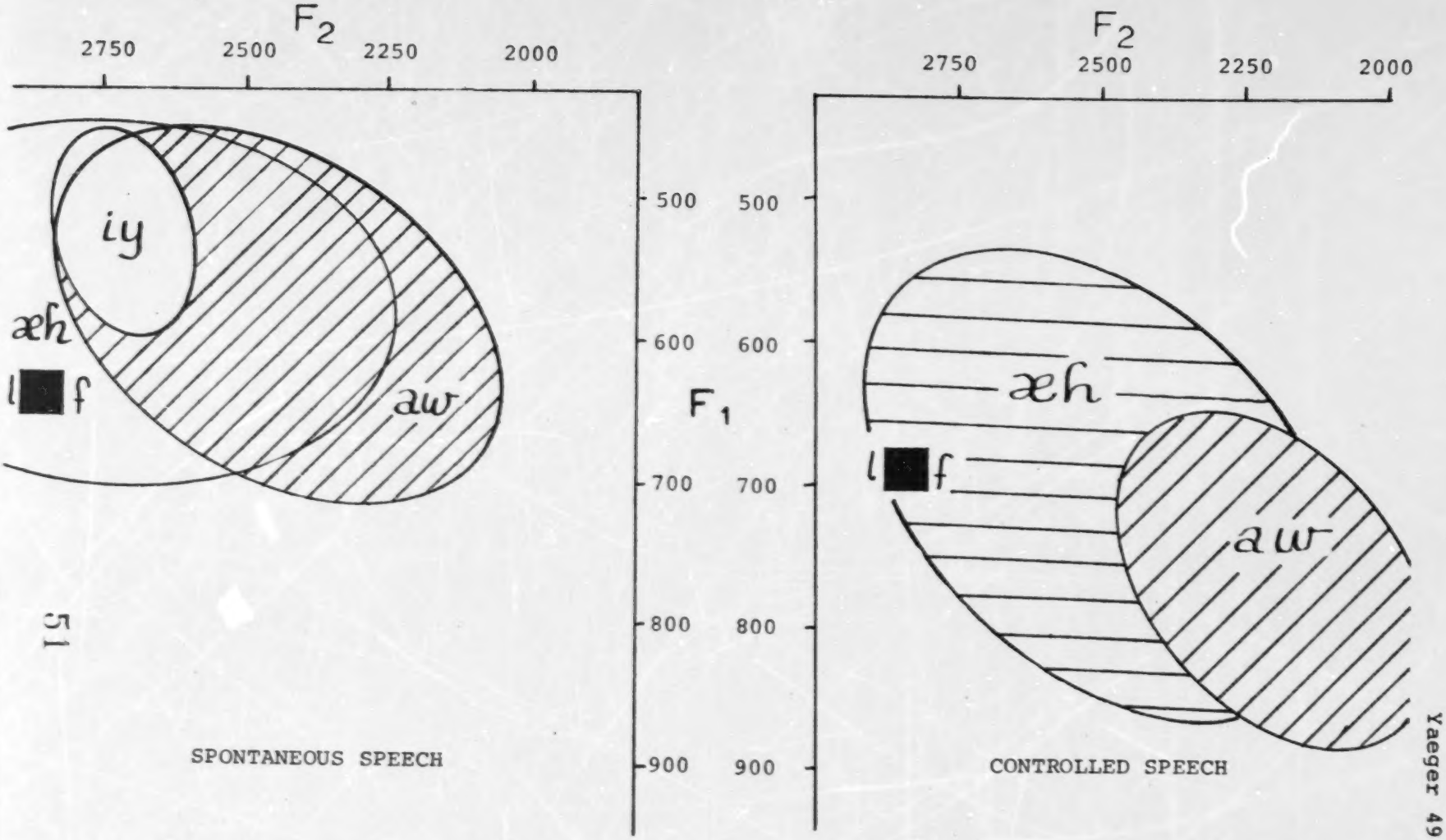


Figure 7. Maintenance of relations between (æh) and (aw) in spontaneous and controlled speech styles of Jane R., 16, Philadelphia.

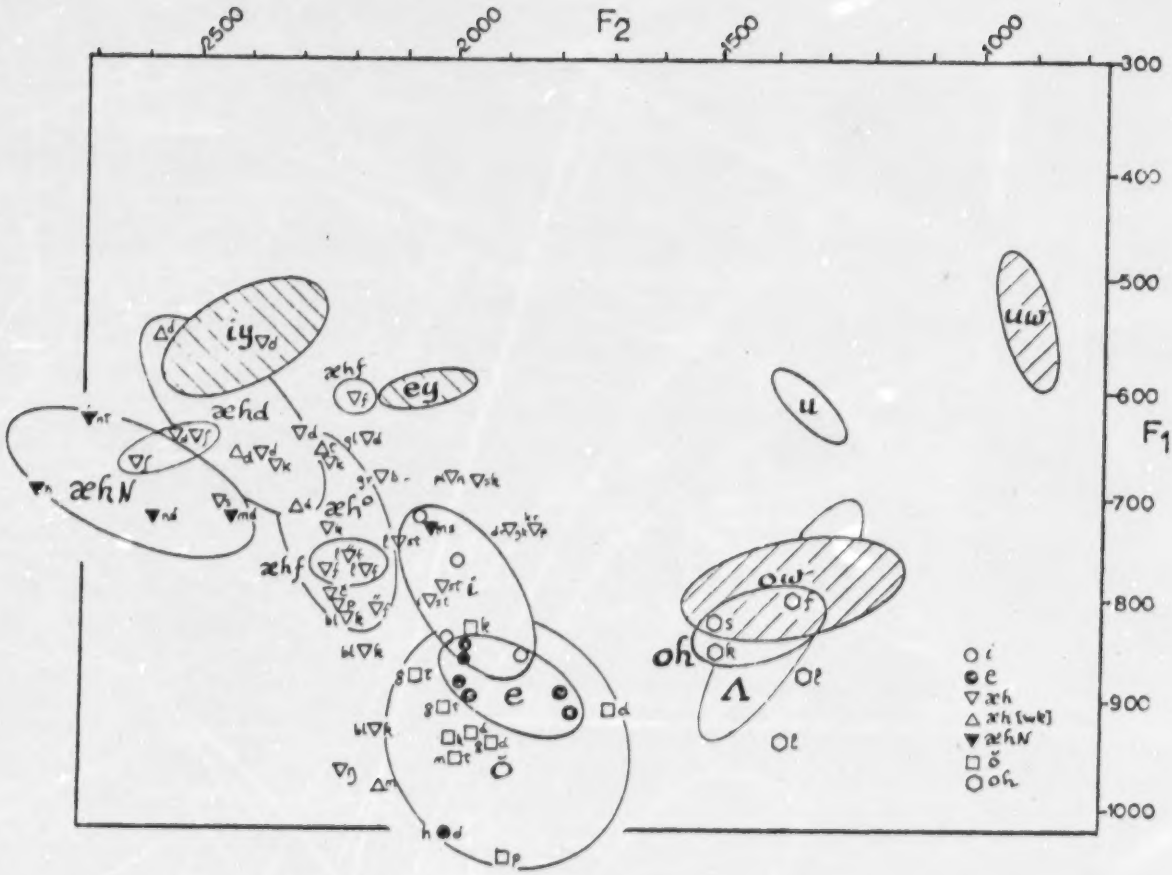


Figure 8a. Phonetic differentiation of (æh) classes in the spontaneous speech of Carol Muehe, 16, Chicago. [Figure 23 from Labov, Yaeger and Steiner 1972].

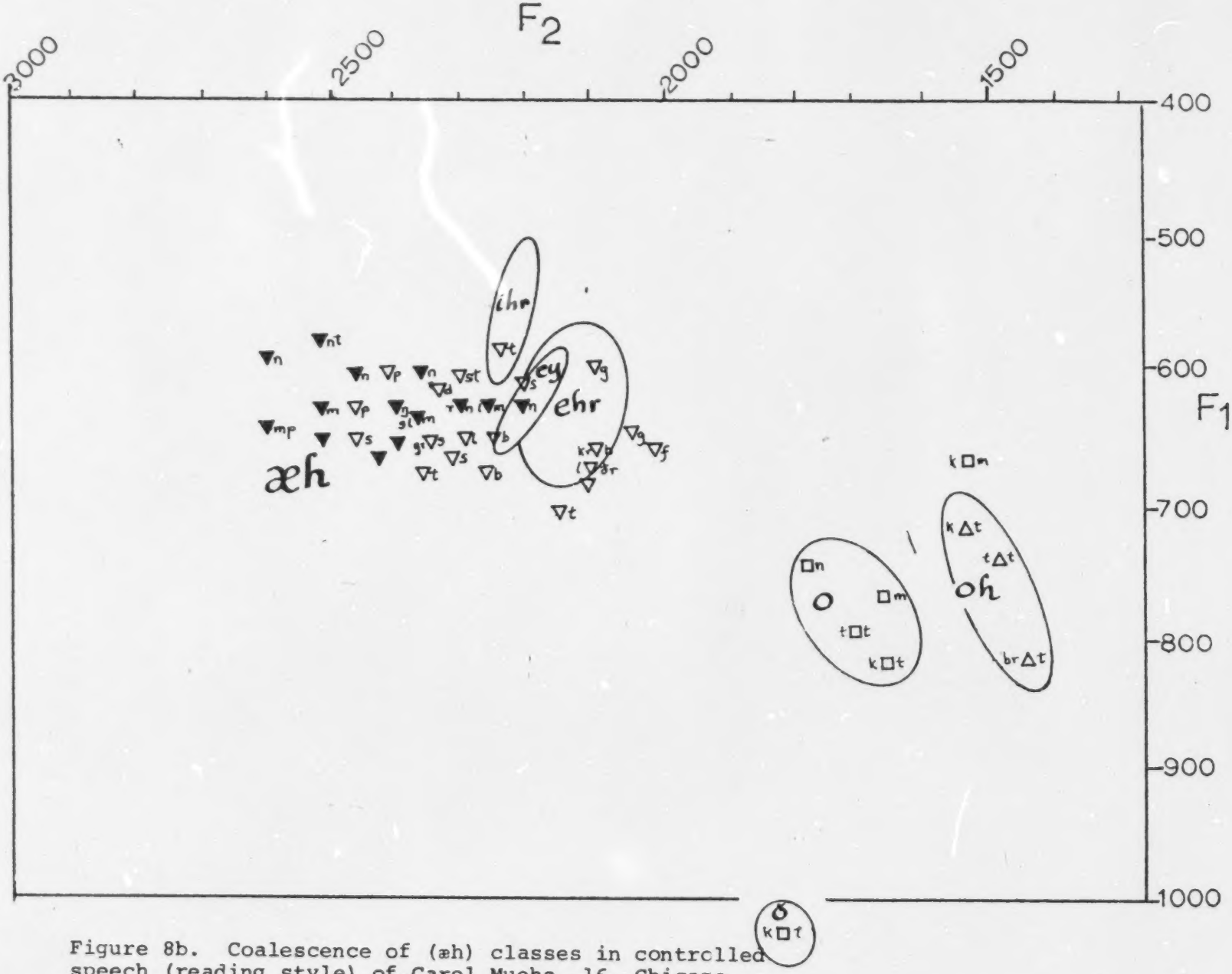


Figure 8b. Coalescence of (æh) classes in controlled speech (reading style) of Carol Muehe, 16, Chicago. [Figure 3-18 from Labov, Yaeger and Steiner 1972].

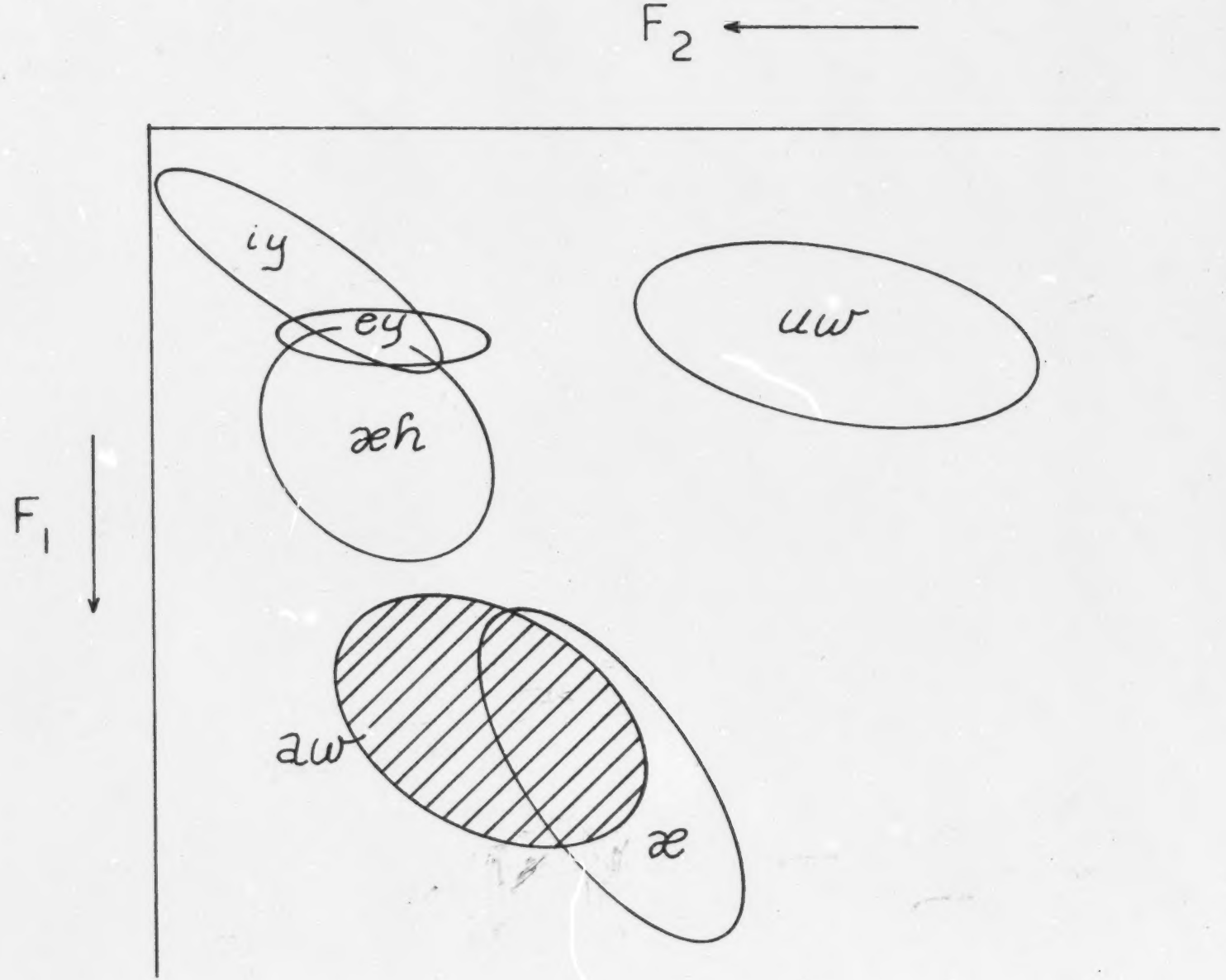


Figure 9. Early stage in the raising of (aw) in spontaneous speech of Bill G., 15, Philadelphia.

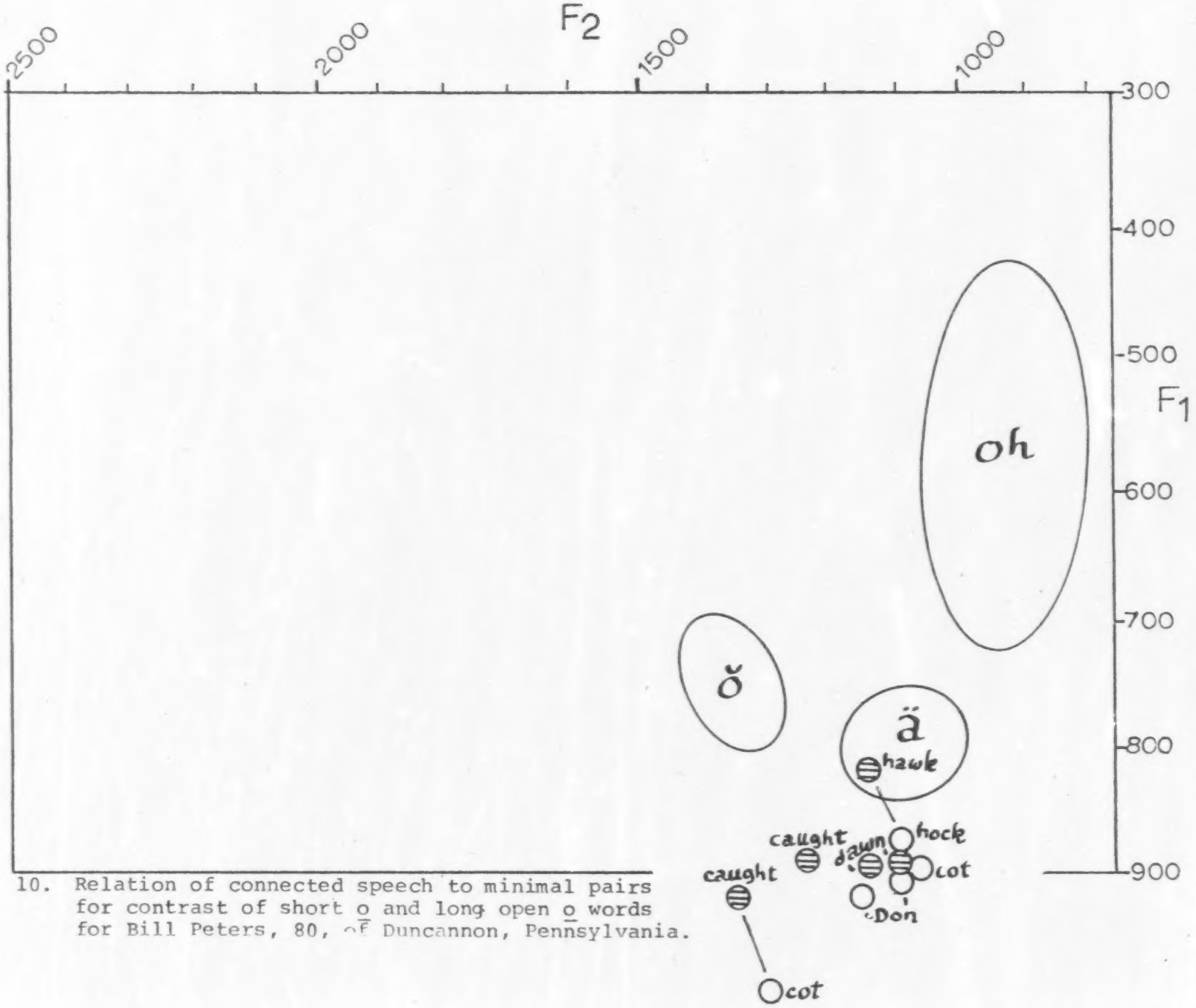


Figure 10. Relation of connected speech to minimal pairs for contrast of short o and long open o words for Bill Peters, 80, of Duncannon, Pennsylvania.

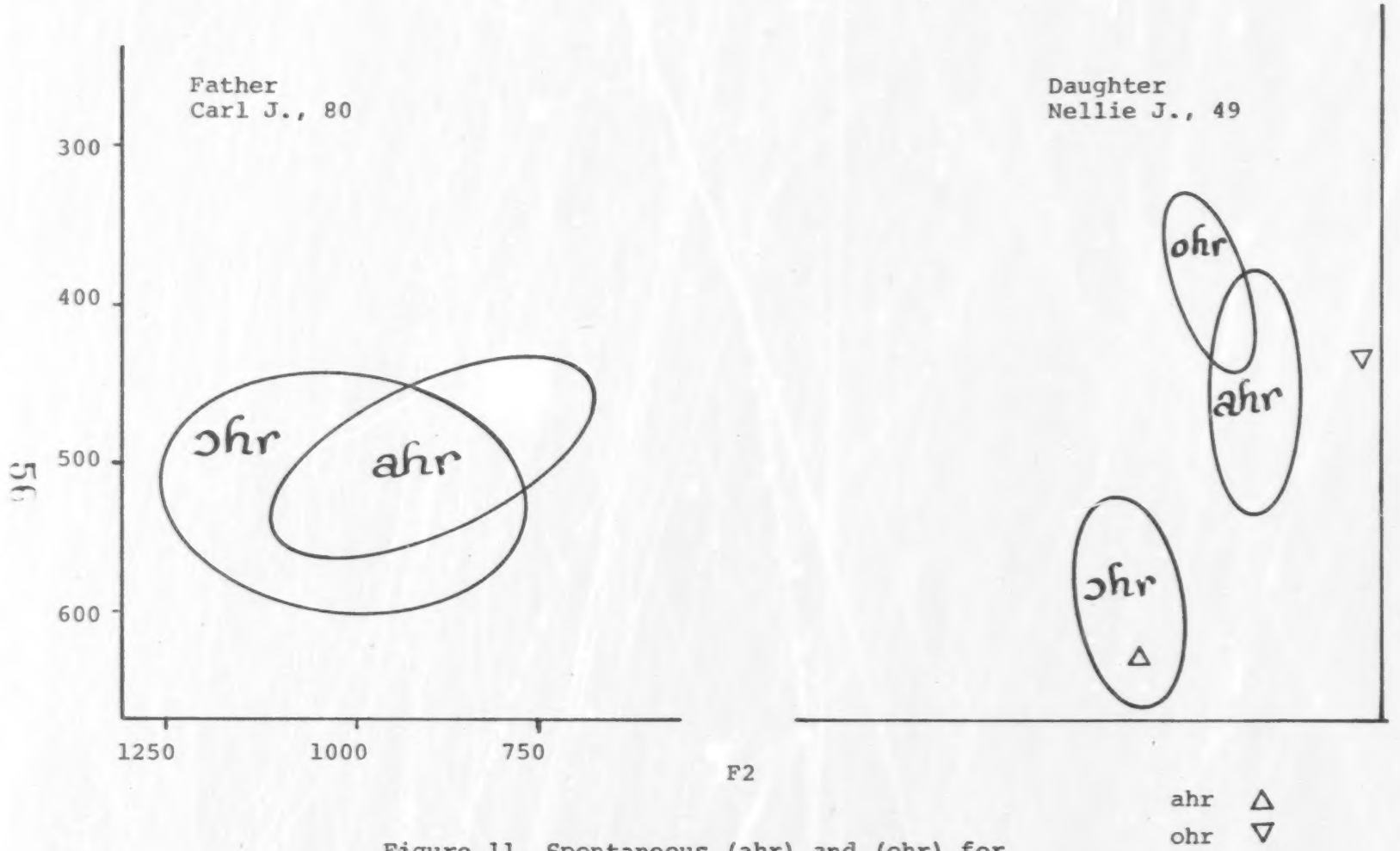


Figure 11. Spontaneous (ahr) and (ohr) for two generations of Southwest Utah (Kanab).

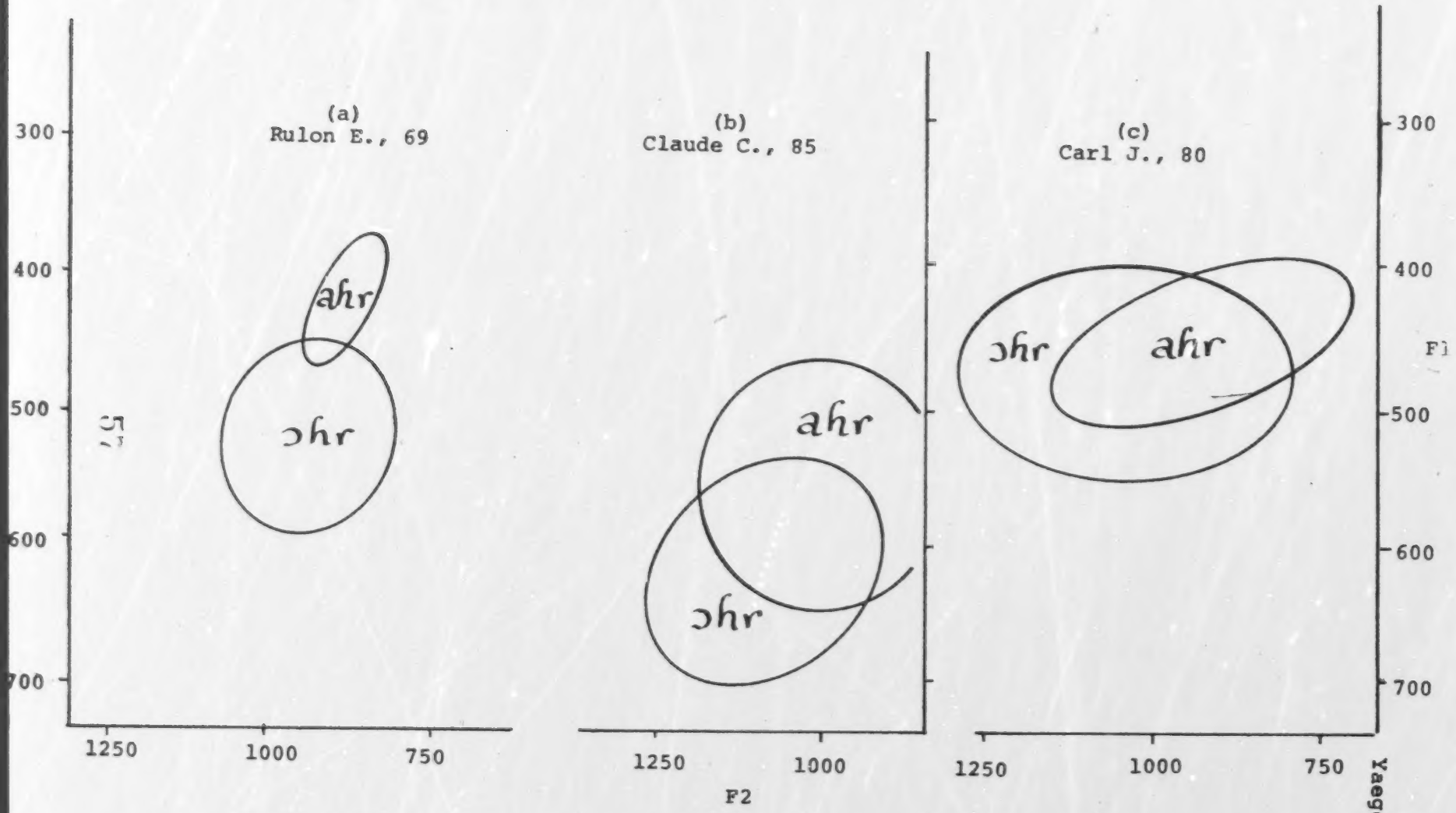


Figure 12. Spontaneous (ahr) and (ohr) for 3 older speakers from Southwest Utah (Kanab).

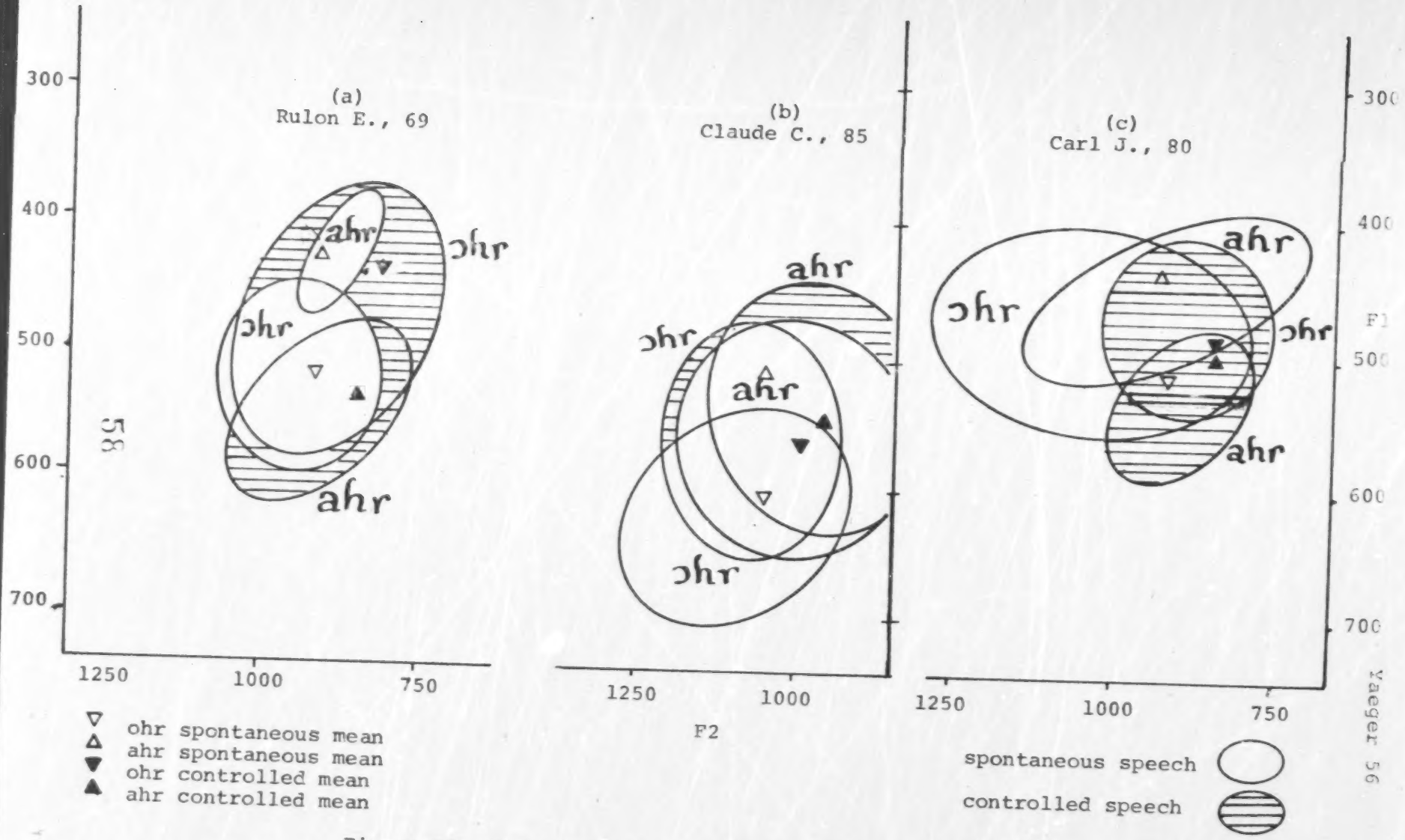


Figure 13. Spontaneous vs. controlled (ahr) and (ohr) for three older speakers from Southwest Utah (Kanab).

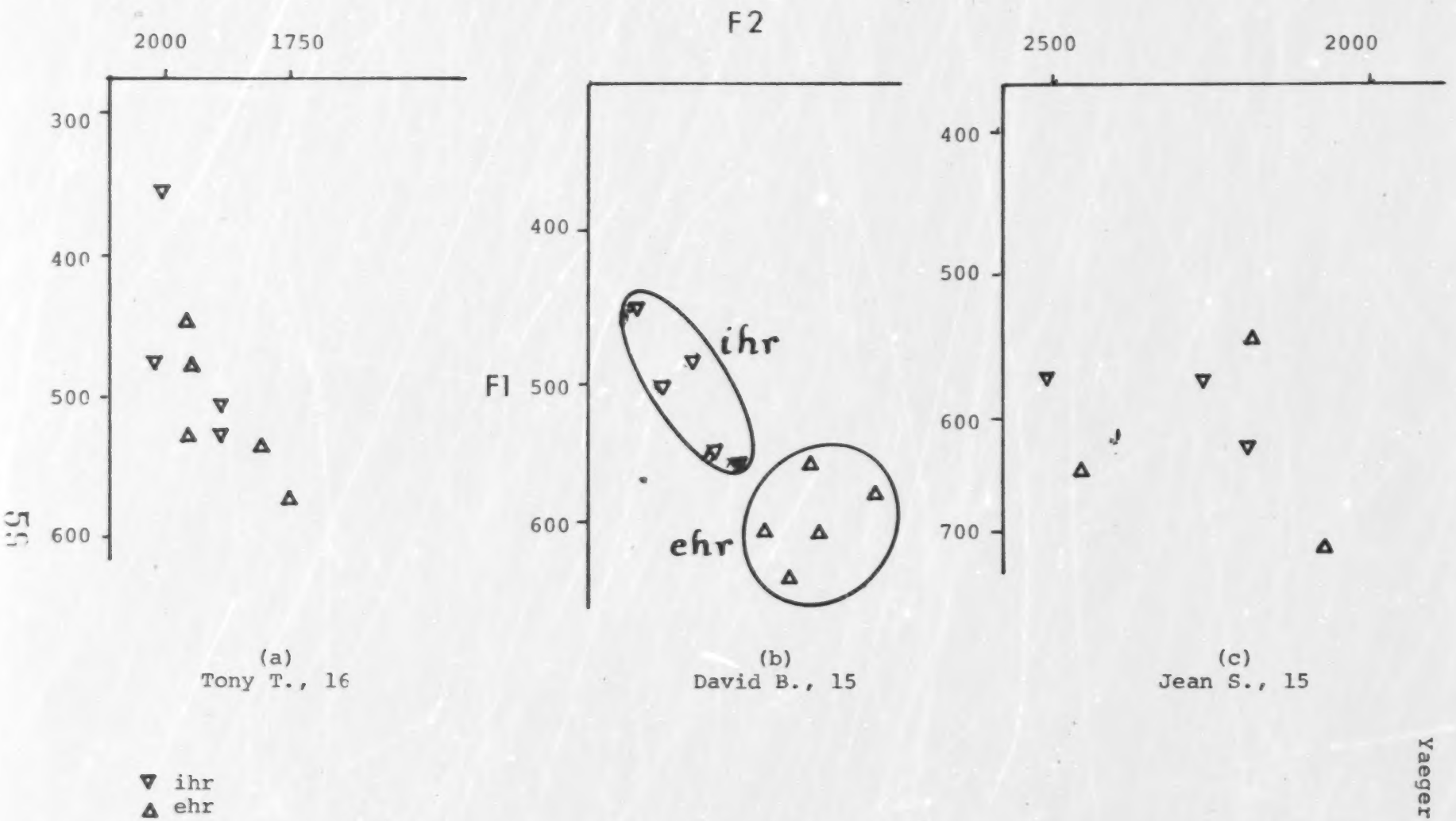


Figure 14. (ihr) and (ehr) in the spontaneous speech of 3 Norwich adolescents.

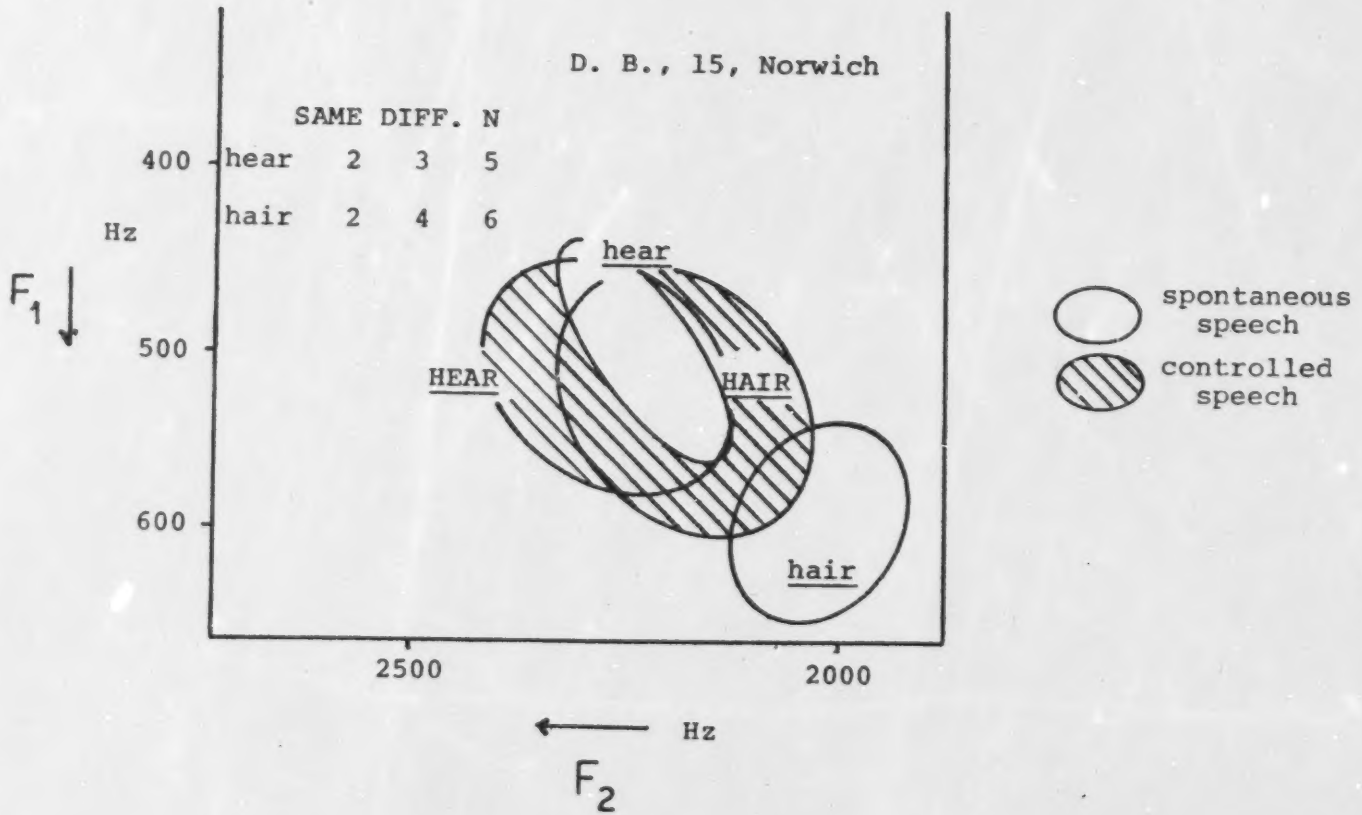
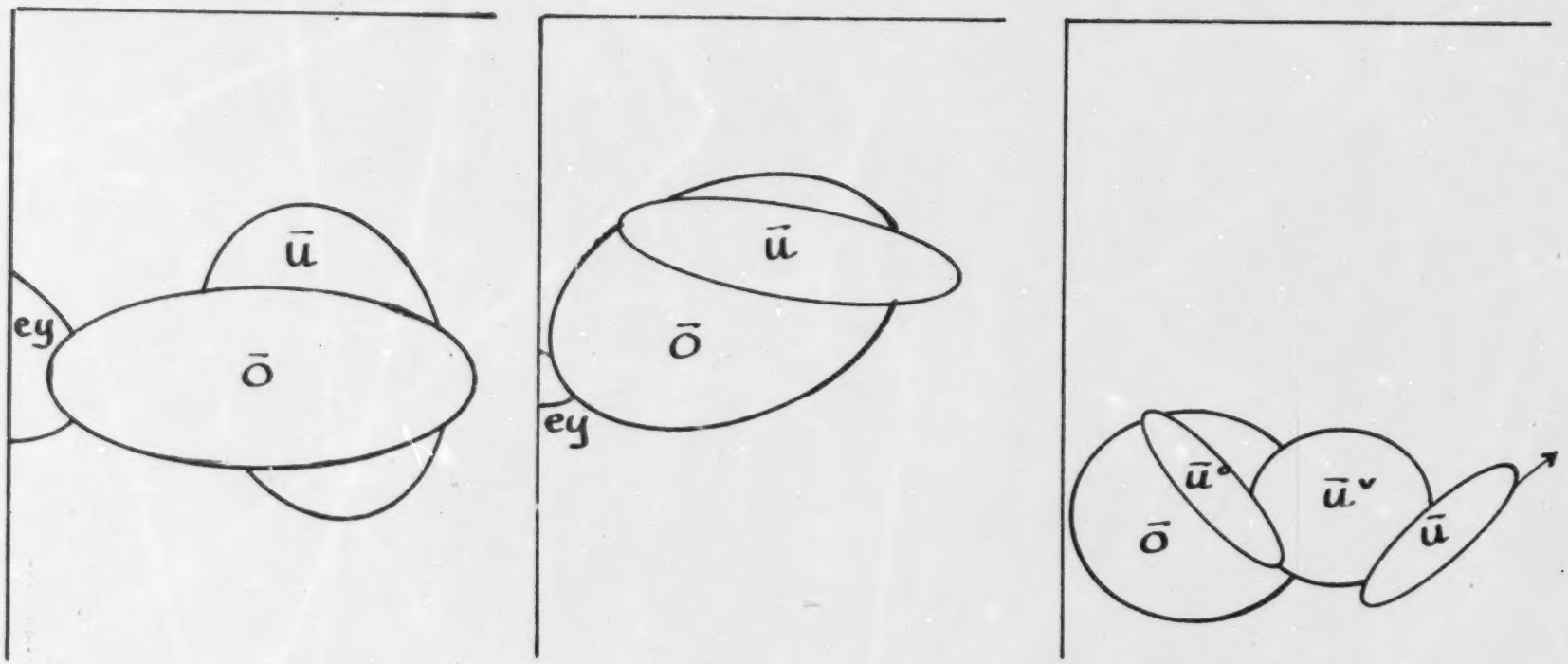


Figure 15. Distribution of (ihr) and (ehr) in spontaneous and controlled production of hear and hair for David B., 15, Norwich, England

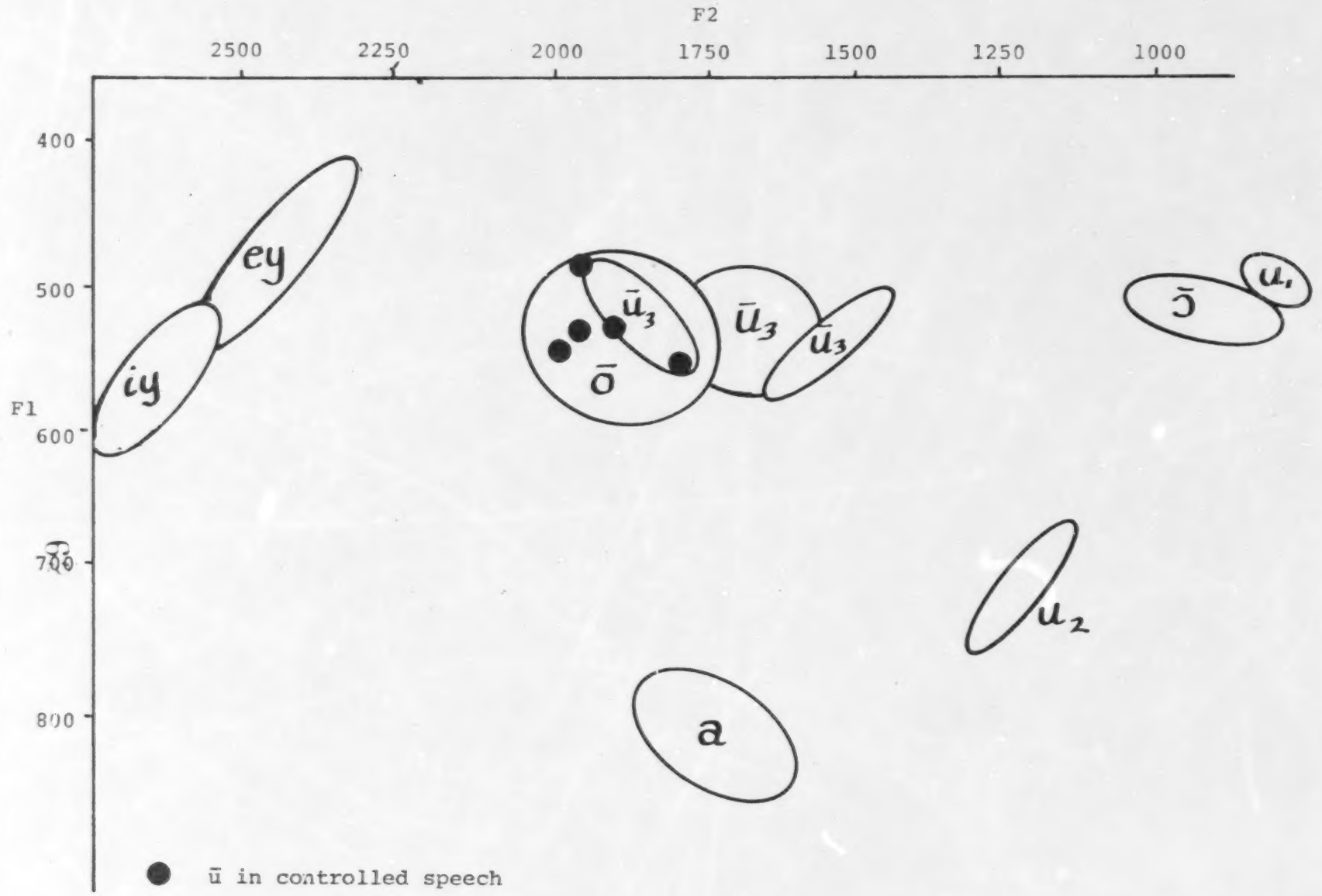


(a)
Mary W., 79

(b)
Jim McN., 21

(c)
Totty R., 14

Figure 16. Distribution of M.E. ō in boot, book, etc. and M.E. ū in house, about, etc. in the spontaneous speech of three Glaswegians.



Yaeger 60

Figure 17. Distribution of (ū) norms in house, about, etc. in spontaneous speech of Totty R., 14, Glasgow.

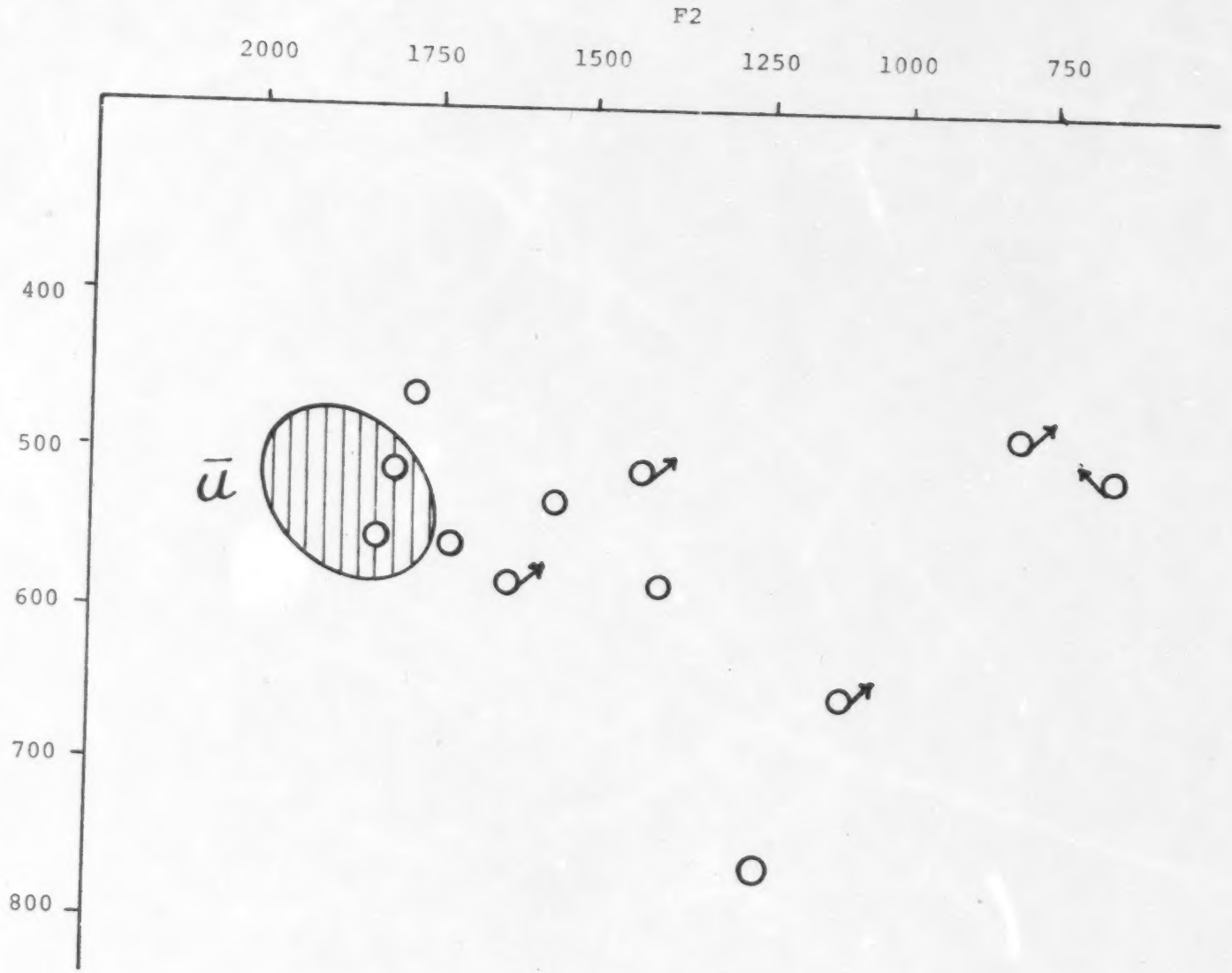


Figure 18. Spontaneous and controlled (ū) for Totty R., 14, Glasgow.